

Tallinna Ülikool
Digitehnoloogiate Instituut

Magistritöö

**RIIGI INFOSÜSTEEMIDE KOOSVÕIME TAGAMINE
ESIMESEKS REGISTRIPÕHISEKS LOENDUSEKS EESTIS**

Juhendaja: Arne Ansper
Koostaja: Diana Beltadze

Autor: „ „ 2016
Juhendaja: „ „ 2016
Instituudi direktor: „ „ 2016

Tallinn 2016

AUTORIDEKLARATSIOON

Deklareerin, et käesolev magistritöö on minu, Diana Beltadze, töö tulemus ja seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on viidatud.

..... (allkiri)

..... (kuupäev)

Lihlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Diana Beltadze (sünnikuupäev: 6.08.1969)

annan Tallinna Ülikoolile tasuta loa (lihlitsentsi) enda loodud teose „Riigi infosüsteemide koosvõime tagamine esimiseks registripõhiseks loenduseks Eestis“, mille juhendaja on Arne Ansper, säilitamiseks ja üldsusele kättesaadavaks tegemiseks Tallinna Ülikooli Akadeemilise Raamatukogu repositooriumis.

Olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.

Kinnitan, et lihlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tallinnas, _____ allkiri ja kuupäev

ANNOTATSIOON

Magistritöö „Riigi infosüsteemide koosvõime tagamine esimeseks registripõhiseks loenduseks Eestis” käsitleb registripõhist loendust ja infosüsteemide koosvõime tagamise vajalikkust. Autor tutvustab koosvõime tagamise ülesannet ja lahendamise võimalusi ning Statistikaameti võtmerolli vajaliku tulemuse saavutamisel.

Töö fookuses on avaliku sektori organisatsioonide infosüsteemide koosvõime protsess esimese registripõhise loenduse korraldamisel Eestis. Eesmärk on selgitada välja takistavad tegurid infosüsteemide koosvõime tagamisel registripõhiseks loenduseks ja leida strateegiad, mis aitaksid olukorda lahendada meetoodiliselt. Tagamõte on rõhutada andmekogude andmekvaliteedi olulisust infosüsteemide koosvõime tagamisel ning selle aktuaalsust.

Horisontaalse koosvõime tasandil on enim olulised registripõhise loenduse tegemisel kasutada standardeid, parandada andmekvaliteeti ja edendada semantilist võimekust. Infosüsteemide koosvõime tagamise eeldus on, et infosüsteemides rakendatakse riigi loodud kindlustavaid süsteeme nt aadressiandmete süsteemi, klassifikaatorite süsteemi jn.

Magistritöö koosneb kaheksast peatükist. Esimene tutvustab infosüsteemide koosvõimealast kirjandust. Teises peatükis avatakse registripõhise loenduse mõiste. Selle paigutamine töö algusesse on vajalik, sest see on tihedasti seotud töö eesmärgiga. Tuuakse välja registripõhise loenduse eelised. Kolmandas peatükis tutvustatakse statistika tegija rolli registripõhise loenduse tegemisel. Neljandas peatükis kirjeldatakse registripõhise loenduse ettevalmistamise ajalugu Eestis. Viiendas peatükis tutvustatakse riigi infosüsteemide seisuga registripõhise loenduse läbiviimiseks. Antakse ülevaade infosüsteemide erinevatest eesmärkidest ja arengutasemest tulenevatest probleemidest. Kuuendas peatükis on kirjeldatud registripõhise loenduse vajadusi ja nõudeid infosüsteemide koosvõimele. See on töö keskne osa, mis uurib, mida on tegelikult vaja koosvõime tagamiseks. Seitsmendas peatükis on esitatud registripõhise loenduse andmehõive esitlusmudelit. On selgitatud, milliseid nõudeid peavad täitma selleks andmekogud ning statistika tegija. Kaheksas peatükk käsitleb vastuseid uurimisülesannetele ja seda, kas hüpoteesid leidsid kinnitust.

Töös on 11 joonist, 8 tabelit ja 10 lisa. Kokku 117 lehekülge.

SISUKORD

ANNOTATSIOON.....	4
SISUKORD.....	5
SISSEJUHATUS.....	8
Töö eesmärk	8
Teema aktuaalsus	8
Probleemi sõnastus	9
Uurimisülesanded.....	9
Uurimismeetod	10
Uurimisobjekt	10
Magistritöö hüpoteesid	10
Magistritöö tulemus	11
Magistritöö struktuur	11
1. INFOSÜSTEEMIDE KOOSVÕIMEALASE KIRJANDUSE LÜHIÜLEVAADE.....	13
2. REGISTRIPÕHINE LOENDUS	19
2.1 Loenduse mõiste	19
2.2 Registripõhise loenduse standardi nõuded	20
2.3 Registripõhise loenduse eelised	22
2.4 Registripõhine loendus maailmas	22
3. STATISTIKAAMETI ROLL REGISTRIPÕHISE LOENDUSE TEGEMISEL	26
4. PROBLEEMI ANALÜÜS.....	29

4.1 Registripõhise loenduse eellugu.....	29
4.2 Registripõhise loenduse ettevalmistamine	35
4.3 Registripõhise loenduse õiguslik raamistik.....	39
5. RIIGIINFOSÜSTEEMID REGISTRIPÕHISEKS LOENDUSEKS	44
5.1 Infosüsteemide võrgustik loenduseks	46
5.2 Infosüsteemide erinevad eesmärgid ja tasemed	48
5.3 Riigiinfosüsteemide koosvõime raamistik registripõhiseks loenduseks.....	50
6. REGISTRIPÕHISE LOENDUSE VAJADUSED JA NÕUDED INFOSÜSTEEMIDE KOOSVÕIMELE	53
6.1 Riigi infosüsteemide loenduseelne seis	53
6.2 Strateegiad registripõhise loenduse infosüsteemide koosvõime tagamiseks	54
6.2.1 Infointegratsioon.....	56
6.2.2 Protsessiintegratsioon	58
7. ANDMEHÕIVE LOENDUSEKS	61
7.1 Andme esitusmudel andmehõiveks	64
8. ARUTELU	68
KOKKUVÕTE.....	74
ETTEPANEKUD	75
RESÜMEE	76
KASUTATUD ALLIKAD	78
KASUTATUD KIRJANDUS.....	83

JONISED JA TABELID.....	94
LISAD.....	96
LISA 1. LOENDUSE TUNNUSED	97
LISA 2. REGISTRITE KASUTUSELE VÕTMINE LOENDUSEKS PÕHJAMAADES	99
LISA 3. SEOTUD ISIKUD JA AADRESSID MAAKONDADE LÕIKES 2011. AASTAL.....	100
LISA 4. KOOSVÕIME TAGAMISE TEEKAART LOENDUSEKS	101
LISA 5. LOENDUSANDMETE ANDMEHÕIVE MEETODITE STRATEEGIA.....	102
LISA 6. PROBLEEMI KIRJELDUSED LOENDUSTUNNUSTE RÜHMES	103
LISA 7. RIIKLIKE INFOSÜSTEEMIDE KOOSVÕIMELISUSE SEIS LOENDUSEKS	104
LISA 8. ZACHMANI RAAMISTIK	106
LISA 9. ANKEETKÜSITLUSE KÜSIMUSED	107
LISA 10. ANKEETKÜSITLUSE VASTUSED	109

SISSEJUHATUS

Magistritöö „Riigi infosüsteemide koosvõime tagamine esimeseks registripõhiseks loenduseks Eestis” fookuses on avaliku sektori organisatsioonide infosüsteemide koosvõime protsess esimese registripõhise loenduse korraldamisel Eestis. Käsitlen töös infosüsteemide koosvõime aspekte, mis horisontaalse koosvõime tasandilt on olulisemad registripõhise loenduse tegemisel: infoallikate integratsioon, standardite kasutamine, andmekvaliteedi parandamine, semantikaalne võimekus jt. Arutlusele ei tule infosüsteemide koosvõime turbealased aspektid, arhitektuurialased kirjeldused, adaptatsioonid, konsolideerimisalased võimekused jn. Töö huviväärsus haarab loenduse tööloiku, mis on piiritletud andmehõive töödega.

Esmakordse registripõhise loenduse korraldamisel Eestis on tarvis teha palju eeltöid infosüsteemide koosvõime tagamiseks, kuluefektiivsus saavutatakse ajalisel plaanis siis, kui on saavutatud infosüsteemide koosvõime ja võimalik korraldada registripõhiselt loendusi sagedamini.

Töö eesmärk

Magistritöö eesmärk on selgitada välja kitsaskohad infosüsteemide koosvõime tagamisel registripõhiseks loenduseks ja leida esile kerkinud probleemidele lahendusi ning uurida välja, millised on võimalused tagada infosüsteemide koosvõime. Töö fookuses on infosüsteemide koosvõime tagamine Eesti avalikus sektoris. Registripõhise loenduse infosüsteemide koosvõime tagamine on koosvõime üks erijuhtum. Selle töö eesmärk on aidata kaasa andmekogumise protsessi parendamisele, muutes veidi andmehõive loogikat. Koostatud andmehõive mudel tugineb unikaalandmete kasutamisele, mis eeldab semantilist võimekust, kvaliteetseid registriandmeid loenduse jaoks ja andmehõivet X-tee andmevahetuskeskkonna kaudu. Magistritöö tagamõte on näidata, kui oluline on andmekogude kvaliteet ning aktuaalsus infointegratsiooni perspektiivis.

Teema aktuaalsus

Uurimisteema omab praktilist väärtust, et teha registriandmetel põhinevat statistikat kiiremini. Infosüsteemide koosvõime registripõhiseks loenduseks on uudne teema, kuid infosüsteemide koosvõime temaatika on varem käsitletud leidnud. See on rakendust loov uuring, mille käigus kaardistatakse institutsionaalsed teadmised ja oskused, mis on vajalikud teatud funktsionaalseid nõudeid rahuldava tulemuse loomiseks registripõhise loenduse korraldamisel Eestis 2020. aastal.

Rahva ja eluruumide loendusi tehakse tavaliselt iga kümne aasta järel ja see on rahvastikuarengu kirjeldamisel kõige autoriteetsem andmeallikas. Järgmist rahva ja eluruumide loendust plaanib Statistikaamet korraldada esimest korda Eesti ajaloos registriandmetele tuginedes. Teemakäsitus pakub lahendusi statistika tegemise ärioloogika parandamiseks, püüab aidata muuta seda efektiivsemaks ja tulusamaks. Kui infosüsteemide loendusaegne koosvõime on saavutatud, viitab see ühiskonna laiemale koosvõimele – registrid saavad korda, andmete kvaliteet paraneb ning info kättesaadavus ja kasutatavus suureneb.

Probleemi sõnastus

Registripõhise loenduse läbiviimiseks tuleb ette valmistada infotehnoloogilisel lahendusel põhinev maksimaalselt automatiseeritud andmekogumisprotsess. Andmeid tuleb koguda registritest ja need peavad vastama statistika tegemise nõuetele. Registriest andmete kogumisele kuluv aeg ei ole aktsepteeritav, andmehõivele kuluvat ajalist ressursi tuleb vähendada. Probleemiks on andmete kogumise viis. Uurimistöö käigus selgus, et andmehõive protsessi optimeerimise takistavateks teguriteks on:

1. andmete formaat, see ei vasta standartidele (nt aadressiandmed);
2. asutused kasutavad erinevaid viise andmevahetuseks;
3. andmekvaliteet on ebaühtlane.

Uurimisülesanded

1. Selgitan välja, missugused on eeldused infosüsteemide koosvõimeks registripõhise loenduse puhul.
2. Selgitan välja, missugust infosüsteemide koosvõime strateegiat kasutada registripõhiseks loenduseks.
3. Parendan loenduseks vajalike andmekogude andmehõive toimimise mehhanismide efektiivsust.
4. Koostan ettepanekud andmehõive optimeerimiseks ja koostöö edendamiseks registripidajatega.

Uurimismeetod

Magistritöös esitan arendusuuringu, mille raames koostasin koosvõime mudeli teiseks registripõhiseks prooviloenduseks 2018. aastal Eestis. Teema avamiseks kasutan andmete analüüsimisel kombineeritud meetodit. Analüüsisin registripõhise loenduse vajadusi, intervjuerisin andmehõive eest vastutavaid spetsialiste andmete edastajate ja vastuvõtjate poolt andmehõive olukorra välja selgitamiseks. Viisin läbi 2016. aastal ankeetküsitluse andmekogupidajatega eesmärgiga selgitada välja andmekogude pidajate poolsed arusaamad koosvõime saavutamisel.

Uurimisobjekt

Uurimisobjekt on riigi infosüsteemide kaudu toimiv andmekogumise protsess. Infosüsteemid peaksid tagama loendusandmete vastuvõtmise ja edastamise.

Magistritöö hüpoteesid

Hüpotees 1: Riigi infosüsteemide koosvõime tekib, kui andmekogude pidajad panustavad võtmeandmete keskele korrashoiule ja võtavad need kasutusele riiklikes andmekogudes.

Hüpotees 2: Loenduse läbiviimiseks tehtud muudatused riigi infosüsteemides on kasutatavad muude ülesannete lahendamiseks.

Magistritöö tulemus

Magistritöö tulemusena on välja selgitatud kitsaskohad infosüsteemide koosvõime tagamisel ja tehtud ettepanekud, kuidas parandada registripõhise loenduse ettevalmistamiseks tehtavat tööd.

Magistritöö struktuur

Magistritöö koosneb kaheksast peatükist.

Sissejuhatus, probleemiseade

Omaette osadena käsitletakse kõiki töö pealkirjas nimetatud teemasid – registripõhist loendust ja selle vajalikkust, riigi infosüsteeme ja nende võrgustumise hetketaset, koosvõime tagamise ülesannet ja lahendamise meetodikat ja statistikategija roll selle koosvõime tegelikul tagamisel ja vajaliku tulemuse saavutamisel.

Esimene peatükk tutvustab infosüsteemide koosvõimelase kirjanduse ülevaadet.

Teine peatükk avab mõiste registripõhine loendus. Selle paigutamine töösse on vajalik, sest see on tihedasti seotud töö eesmärgiga, milleks registripõhist loendust on tarvis. Toon välja registripõhise loenduse eelised, mis välja toodud registripõhise loenduse kogemusest maailmas.

Kolmandas peatükis tutvustatakse statistika tegija rolli registripõhise loenduse tegemisel.

Neljandas peatükis kirjeldatakse registripõhise loenduse eellugu Eestis, mis toob välja probleemi olemuse.

Viiendas peatükis antakse ülevaade riigi infosüsteemide seisust registripõhise loenduse läbiviimiseks. Esitatakse infosüsteemide erinevatest arengutasemetest tulenevaid probleeme.

Kuuendas peatükis on kirjeldatud registripõhise loenduse vajadused ja nõuded infosüsteemide koosvõimele. See on töö keskne osa – mida on tegelikult vaja koosvõime tagamiseks?

Seitsmenda peatükis on esitatud andmehõive esitlusmudel. On jagatud selgitusi, milliseid nõudeid peavad täitma selleks andmekogud ning statistika tegija.

Kaheksas peatükk, arutelu, käsitleb vastuseid uurimisülesannetele ja kas hüpoteesid leidsid kinnitust.

Kokkuvõte käsitleb koosvõime tagamist. Viitan püstitatud tööhüpoteesidele, kas need said tõestatud.

1. INFOSÜSTEEMIDE KOOSVÕIMEALASE KIRJANDUSE LÜHIÜLEVAADE

Käsitlusi infosüsteemide koosvõimest on avaldatud mitmes võtmes ning loodud teoreetilisi mudeleid ennustamaks infosüsteemide võimekust organisatsioonide ja/või võrgustike tasandil. Kirjanduse ülevaates keskendun järgmistele teemadele, mis seonduvad infosüsteemide koosvõime loomise vajadusega loenduseks:

1. infosüsteemide koosvõime aspektid protsesside juhtimise tasandil, strateegiad, raamistikud;
2. infointegratsiooni käsitlused seoses andmehõivega.

Koosvõime temaatika kesksemad valdkonnad 21. sajandi uurimistöodes võib tinglikult jagada kahte suuremasse rühma. Esimese rühma peateemad on raamistikud, standardid ja võrgustikud, teise rühma peateemad jagunevad koosvõime käsitluse järgi: koosvõime käsitlused juhtimise tasandil (Hagedoorn, 1993; Katz jt, 1985; Kaufmann jt, 2000; Ducq jt, 2004) ja infosüsteemide koosvõime teemad (Fichman, 1999; Zhao jt, 2014). Metodoloogilisi ja teoreetilisi ülevaateid teemakäsitlusest on kirjutatud enamasti semantika ja süntaktika valdkonnas (Ouksel jt, 1999; Sheth, 1999), kuid infosüsteemide koosvõime tagamise meetodika on leidnud uuritavas kirjanduses vähest kajastust v.a mõningad erandid, nt arutelud organisatsiooni arhitektuuri (Chiang jt, 1984) ja talituspidevuse tagamise teemadel (Bhatt jt, 2005).

Infosüsteemide koosvõime uurimustes on valdav retrospektiivne vaade, perspektiivseid uuringuid on infosüsteemide koosvõimest tehtud seoses arendusuuringute, kognitiivse ja ratsionaalse valiku või juhtimisteooriate valguses, matemaatilisi koosvõime mudeleid on loodud semantilise (Ouksel, 1992) koosvõime ning võrgustiku teemadel (Ducq jt, 2004; Panetto jt, 2012). Tehnilise integratsiooni aspekte on käsitletud seoses erinevate platvormidega, võrguseadmete ühendustega, kommunikatsiooni protokollidega jne. Olulist tähelepanu on uurijad osutanud andmete formaadi ja kvaliteedi temaatikale.

Peamistest lähenemistest infosüsteemide koosvõime temaatikale uuritud kirjanduses perioodil 1990-2015 võib jagada kolme rühma:

1. koosvõime ulatuse käsitlused: raamistikud, standardid, võrgustikud;
2. koosvõime käsitlused juhtimise tasandil: organisatsiooni võrgustiku vormid, koosvõime regionaalsed tasandid, äriprotsessid, võrgustike juhtimine koostöö vormide dünaamika jt;
3. infosüsteemide koosvõime teemad: IT arhitektuur, tarkvara, süntaks, semantika, andmete edastuse protokollid jt.

IT-süsteemide võimekust seostatakse IT-arhitektuuri valdkonnaga, millest tulenevalt on uuritud infosüsteemide suutlikkust vahetada infot, töödelda edastatud andmeid ja seostada neid äriotstarbeliste andmetega (Chiang jt 1984). Infosüsteemid peavad avalikus sektoris tagama turvalise andmete kasutuse. Igal infosüsteemil on otstarve. Kui on vajadus seda ümber korraldada, tuleneb see reeglina muutusest äriprotsessides või on seotud otsustusprotsessidega (Ducq jt, 2004: 193-203). Uurijad on märkinud, et edukamad on need lahendused, mis integreerivad „tagasüsteemide” (back-end) funktsioone (Zhu jt, 2005: 61-84). Rai ja Tang on arvamusel, et infosüsteemide koosvõime oleneb kahest protsessist: IT-integratsioonist ja IT-rekonfiguratsioonist (Rai jt, 2010: 516-542). Rai ja jt on tuvastanud tugeva IT-infrastruktuuri seose asutuse väärtusahela protsessidega (Rai jt, 2012: 233-262).

Koosvõime raamistike teema on maailmas enim käsitlust leidnud käesoleval sajandil. Šveitsi teadlased Legner ja Wende on märkinud, et koosvõime kätkeb endas organisatsioonilist ja funktsioneerivat võimekust teha koostööd ja luua väärtusi, selle aluseks on võimekus jagada infot äripartnerite vahel ja luua koostöövõrgustikke (Legner ja Wende, 2006: 1-13). Wigand ja Österle leiavad, et äriline koosvõime on koostöövõimekus (Wigand, 1997; Österle, 2001). Koordinatsiooni teooriast lähtuvalt on nad uurinud organisatsioone ja esitanud eksisteeriva koosvõime problemaatika äritegevuse tasandil. Teemad, mis tihti üles kerkivad, on järgmised: defineerida koostööeesmärgid ja panna need rangesse vormi (lepingud ja SLA-d), organisatsioonidevahelised äriprotsessid, tehnoloogiate ja platvormide valik, ühendada organisatsioonidevahelised infosüsteemid. Eelpool mainitud autorid on jõudnud järeldusteni, et infosüsteemide võimekuse tagamine võimaldab organisatsioonidel liituda uute partneritega ilma täiendava kuluta (*Ibid*).

Legner ja Wende annavad ülevaate maailma suurematest levinumatest koosvõime raamistikest nt IDABS (*European Interoperability Framework*), EIF (*European Interoperability Framework*), e-GIF (*E-Government Interoperability*), LICI (*the Levels of*

Information Systems Interoperability framework), EICTA (*European Information and Commnucation Tehnolocy Industry Assosatsion*) jt. Koosvõime raamistikud on pidevas muutumises, mistõttu soovivad autorid lähtuda juhuslikkuse teooriast ärialase jm koosvõime analüüsi läbiviimisel ning koosvõime mudelite koostamisel organisatsioonides soovivad autorid võtta aluseks EFQM raamistiku (Legner jt, 2006: 5).

Eesti uurijatest on infosüsteemide koosvõimet enim käsitlenud Uuno Vallner. Ta on pakkunud koosvõime raamistiku definitsiooni. Vallneri järgi on koosvõime disparaatsete ja erilaadsete organisatsioonide võime suhelda vastastikku kasulike ja kokkulepitud ühiste eesmärkide saavutamiseks, vahetades organisatsioonide vahel infot ja teadmisi (Vallner, 2009). Küsimus infosüsteemide koosvõimest tagamisest on tihedalt seotud nende omadustega, mis tagavad jätkusuutlikkuse. Võimekust aitavad tagada (Zhao jt, 2014: 269-298):

- reeglid,
- standardid,
- protseduurid,
- meetodid,
- ühtsed või ühildatavad infosüsteemid,
- koostöö võrgustikud.

Arvestama peab infosüsteemi eluiga ning sõltuvusi teistest süsteemist. Mc Ginnes Simon toob välja, et kontseptuaalne sõltuvus on peaaegu universaalne infosüsteemi disain, mis toob endaga kaasa soovimatuid negatiivseid tagajärgi, sealhulgas süsteemi jäikust ja suuremaid hoolduskulusid (Mc Ginnes jt, 2015: 33-50). Tema väite kohaselt tuleks infosüsteemi osi kavandada sõltumata kontseptsioonist ja toota kohaldatavaid infosüsteeme, et tagada parem koosvõime. Petter Stacie on uurinud infosüsteemide edukuse dimensioone (Stacie jt, 2013: 7-62):

- süsteemi kvaliteeti,
- informatsiooni kvaliteeti,
- kasutust,
- kasutaja rahulolu,
- individuaalset- ja organisatoorset mõju.

Ta on andnud kriitilise hinnangu DeLone'i ja McLeani loodud teoreetilisele mudelile, mis ennustab infosüsteemide edu. Tema tunnetusteoreetiline seisukoht on, et infosüsteemi kvaliteedi ettekuulutajad on süsteemikasutajate karakteristikud:

- suhtumine tehnoloogiasse,
- kogemus tehnoloogiaga,
- enesetõhusus.

Muuhulgas viitavad autori uurimistulemused infosüsteemide koosvõime seosele organisatsiooni juhtimisega. Ta väidab, et organisatsiooni strateegia, mis kasutab uuenduslikku infotehnoloogiat, loob tõenäoliselt paremaid äritulemusi. (*Ibid*)

Kuigi infosüsteemi peamine eesmärk on toota ja tagada kvaliteetset informatsiooni, on mõned uuringud avastanud tegureid, mis oluliselt mõjutavad informatsiooni kvaliteeti. Mainitud DeLone ning McLean jõudsid järeldusele, et infosüsteemide edu tagab mitmemõõtmeline konstruktsioon, mis koosneb süsteemi ja teabe kvaliteedist, kasutamise kogemusest, kasutajate rahulolust, individuaalsetest ja organisatsiooni mõjutavatest teguritest. Samu aspekte infosüsteemide edu tagamise kohta märkisid Shannon ja Weaver juba 1949. aastal (Shannon jt, 1949). Wixom ja Todd väidavad, et tänapäevases muutuv keskkonnas ei ennusta kasutajate rahulolu eriti süsteemi kasutamise edukust, pigem tuleks edu puhul arvesse võtta tehnoloogia tunnustamist ning selle kasutuselevõttu (Wixom ja Todd, 2005). Kui veel eelmise sajandi lõpus Ouksel ja Igbal olid arvamisel, et semantilise koosvõime tagamine on vaevaline protsess, siis 21. sajandi tarkvara arendustööde uurimustes etendab semantilise koosvõime tagamine olulist rolli (Ouksel jt, 1999: 1-29). Semantika raamistik esitab nõuded arendada keskkonda, mis võib soodustada sõltumatu infoallika ebakindlust (*Ibid*).

P. Tallon on oma uurimuses esile tõstnud aspekti, et viimastel aastatel on infosüsteemide haldajad hakanud mõistma, et info on strateegiline ressurss ja andmete eksponentsiaalne kasv seoses alg (toor)andmete kogumisega ja säilitamisega kogu organisatsioonis on äärmiselt kiiresti kasvav. (Tallon jt, 2013: 141-178.) Ta uuris, miks mõnes organisatsioonis saadakse infosüsteemide andmetest kiiret teavet organisatoorse juhtimise parandamiseks ja samas on praktikaid organisatsioonidest, kus teabe kasutamist võidakse edasi lükata või piirata infovalitsemise tavade tõttu. Uuringu tulemusena märgib autor, et teabe kasutamist

juhtimistavades on võimalik avada väärtuste võtmes. Alati peab küsima, miks ja missuguseid andmeid vaja on, et olla oma äris edukas (*Ibid*). Mitu autorit (Hirschheim ja Klein) rõhutavad uut perspektiivi infosüsteemide koosvõime arengus, mis on seotud elustiili muutustega või tehnoloogia arenguga nn füüsilises keskkonnas (Hirschheim jt, 2012) (nt Eestis ID-kaardiga allkirjastamine alates 2000. aastast või samast perioodist andmete ristikasutamine X-tee kaudu).

Koosvõime vajaduse temaatika on üldjuhul seotud kahe põhjustajaga:

- IKT süsteemidel põhinev andmevahetus (Riigi infosüsteemide koosvõime, 2011).;
- Infosüsteemide heterogeensus (Ouksel, 1999).

Viimasel kümnendil on infosüsteemide kaudu äritegevuse automatiseerimine põhjustanud andmete kuhjumise ehk nn andmete revolutsiooni (Kitchin, 2014). Aktuaalne on luua koosvõimelisi infosüsteeme, mis aitavad vajalikku infot koguda, kaevata, töödelda ja organiseeritult esitleda. Informatsiooni hoidmine, töötlemine ning eelkõige selle edastamine IT-süsteemide kaudu on pannud ühiskonna olukorda, kus tuleb leida uusi organiseeritud infotöötamise lahendusi nt pilvelahendused (*Ibid*). Reflekteerides uurimistulemusi võib infosüsteemide koosvõimet pidada infoühiskonna arengutaseme näitajaks, mida saab defineerida võimekusena ja oskusena vahetada ning kasutada ajas muutuvat informatsiooni.

Riiklikes strateegiates käsitletakse infoühiskonda kui valitsevat elukorraldust, kus enamik inimkonna loodud väärtusi on kätketud teabesse, mida hoitakse, teisendatakse ja edastatakse universaalsel digitaalsel kujul (Praust, 1998). Selleks loob võimalused üleüldine andmeedastusvõrk, millele juurdepääs on tagatud kõigile ühiskonna liikmetele (Majandus- ja Kommunikatsiooniministeerium, 2014). Infoühiskonna üks võtmeküsimusi on tagada organisatsioonide ja infosüsteemide koosvõime. Esseist Andres Langemets on eelmise sajandi 90ndatel määratlenud infoühiskonna sisu läbi selle ekspansiivse loomuse, mida iseloomustab leviulatus ja hõlpsus (Langemets 1998). Eesti tahab olla edukas e-riik. Uutel võimekustel põhinevad töökeskkonnad aitavad kokku hoida halduskoormust ja vähendada tööde maksumust. Eirata ei tohiks asjaolu, et pole olemas ainult digitaalset majandust ega virtuaalset ühiskonda (Sassen, 2002: 1-24). Eestis elades on meie jaoks oluline, et saame kasutada infoühiskonna hüvesid eeldusel, et selleks on vajalikud tehnilised ressursid. Nii üksikisiku kui riigi seisukohalt on ja jääb loenduste jaoks kõige olulisemaks küsimuseks turvalisus ning selle privaatsuse aspekt (Beltadze, 2015). Riik hoiab meie kohta andmekogudes tohutul hulgal

privaatset ja delikaatset informatsiooni. Tehnoloogilised rakendused peavad võimaldama andmekaitset kõige kõrgemal tasemel.

Kokkuvõtvalt võib märkida uuritava teema mitmekesisist, mitmetahulist käsitlust infoühiskonna olemuse, arengu ja koosvõimete kohta. Üldjuhul käsitletakse infosüsteemide koosvõimet seoses infointegratsiooniga rõhutamaks info ja infokäitluse mahtude suhtelist suurenemist infoühiskonnas (Klischevski, 2004). Koosvõime raamistikud kehtestavad reeglid infosüsteemidele ja nende arendamisele, mida peavad järgima infosüsteemide arendajad (Riigi infosüsteemide raamistik, 2011). Taotlused koosvõimeks annavad võimaluse luua efektiivsemaid ja ökonoomsemaid lahendusi. Uued tehnoloogiad jõuavad tarbijani kiiresti ning informatsiooni mahud ja tähtsus kasvavad. Standarditel ja avatud platvormidel põhinev koosvõime on eeldus saavutamaks Euroopa digitaalarengu tegevuskava „Euroopa 2020” eesmärgi (Majandus ja Kommunikatsiooniministeerium, 2014). Seega pole koosvõime ainult infotehnoloogiline küsimus, vaid käsitleb infoühiskonna erinevaid tahke laiemalt. Hallata kasvavat infomahtu, edastada seda kiirelt ja kadudeta ning tagada samas selle terviklikkus, käideldavus ja kvaliteet on saanud uute tehnoloogiate väljakutseks ja koosvõime näitajaks (Vallner, 2009).

2. REGISTRIPÕHINE LOENDUS

2.1 Loenduse mõiste

Antud peatükis avan registripõhise loenduse mõiste, sest see on olulisel määral seotud töö eesmärgiga. Toon välja registripõhise loenduse eelised. Viitan mõnedele näidetele registripõhise loenduse kohta maailmas.

Rahvaloendus on protsess, kus riigis kasutades selleks ette valmistatud isikuid ja metoodikat, viiakse läbi oma haldusalas olevate elanike loendus (ÜRO, 2006: 6-9). Rahvaloendust korraldatakse elanike arvu, soo, rahvuse, paiknemise, majandusliku ja sotsiaalse jaotumuse ja muu teadasaamiseks. Rahvaloenduse protsess koosneb andmete programmilisest kogumisest, analüüsimisest, üldistamisest ja avalikustamisest (ÜRO, 2006: 3-6). Üldiselt on rahvaloendus maailmas ankeedipõhine küsitlus, mis haarab kogu riigi püsielanikkonda, st selles riigis alaliselt elavaid kodanikke ja pikaajalise elamislooga võõrriikide kodanikke (ÜRO statistika division, 2011). Standardsetl ühendatakse rahvaloendusi eluruumide loendusega, et saada teada elanikkonna elutingimusi.

Rahvaloendusi viiakse läbi kõigis riikides, kokkuleppeliselt (vähemalt) iga kümne aasta järel, järjekordne voor toimub maailmas aastail 2020/2021 (ÜRO, 2015). Rahvaloendusi koordineerib ÜRO, Euroopas Euroopa Liidus Eurostat. Rahvaloendus on riigile kohustuslik ettevõtmine (EL määrus nr 763/2008). Euroopa Liidu liikmesriigid saavad loendusstatistika tegemisel toetuda mitmesugustele loenduse viisidele (*Ibid*):

- tavapärasele loendusele;
- registripõhisele loendusele;
- tavapärase loenduse ja valikvaatluse kombinatsioonile;
- registripõhise loenduse ja valikvaatluse kombinatsioonile;
- registripõhise loenduse ja tavapärase loenduse kombinatsioonile;
- asjakohasele uuringule vahelduvate valimitega (jadaloendus).

Eurostat, kohustab liikmesriike tegema rahvaloenduse ettenähtud aastal ja määrab täpse kohustuslike väljundtunnuste loendi ja nende esitamise vormi, tagamaks kõigi liikmesriikide andmete võrreldavuse (*Ibid*). Rahva ja eluruumide loendus on kõikne uuring, mida iseloomustab ühtne programm ja täpselt määratud loendushetk. Järgmine loendushetk on planeeritud 31.detsembril 2020. aastal. Loendusmetoodika valik on riigil vaba, kuid see peab rahuldama kvaliteedinõudeid.

Registripõhine loendus on loendusmeetod, mille kohaselt andmete kogumine toimub registritest (ÜRO, 2007: 43). Registripõhist loendust saab korraldada, kui on ühiskonnas tagatud eeldused selleks (ÜRO, 2007: 7-9):

1. andmete olemasolu registrites;
2. andmeid on võimalik seostada;
3. andmekasutus on tagatud õiguslike regulatsioonidega.

Registripõhise rahva- ja eluruumide loenduse läbiviimiseks on tarvis, et riigis oleksid olemas täielikud, pidevalt aktualiseeritavad ja auditeeritavad registrid, mis kajastavad rahvaloenduse põhiküsimusi (Beltadze, 2015). Lisaks sellele peavad erinevad registrid olema ühitatavad (lingitavad), selleks peavad nad omama ühiseid indikaatoreid (isikukoodid, aadressid).

2.2 Registripõhise loenduse standardi nõuded

Statistikaamet esitas nõuded andmekogude pidajatele, mis tuleb täita registripõhise loenduse alguseks (VVLK, 16.01.2015):

- 1) kaetus populatsiooni suhtes peab olema vähemalt 97%;
- 2) infot loenduseks vajalike tunnuste osas aktualiseeritakse regulaarselt ning see dokumenteeritakse;
- 3) vormikohane identifikaator on olemas vähemalt 95% kirjetest;
- 4) loenduse jaoks vajalike tunnuste kaetus on vähemalt 95%;
- 5) loenduse jaoks vajalike tunnuste väärtuses ei esine olulisi või tehnilisi vigu rohkem kui 1% juhtudest.

Loendus on reglementeeritud statistikatöö, mille puhul on soovituslik järgida järgmisi standardeid:

- 1) riikliku statistika tegemise põhiprotsessi mudel;
- 2) statistilise info mudel;
- 3) statistikategemise arhitektuur;
- 4) statistiliste andmete ja metaandmete vahetamise standard.

Klassifikaatorite osas peab Statistikaamet kooskõlastama riigi infosüsteemi kuuluvates andmekogudes kasutatavad klassifikaatorid ja peab kasutama loendusvoorus kokkulepitud klassifikaatoreid (ÜRO, 2008: 35-129). Klassifikaatorite ja terminoloogia kooskõlastamine riiklikes andmekogudes on registripõhise loenduse ettevalmistamisel üks kesksemaid tegevusi.

Vastavalt riikliku statistika seadusele § 50 (2) on Statistikaametil kohustus hinnata andmekogude kvaliteeti (Riiklik statistika seadus, 2010). Standarditele vastav metaandmete kirjeldamine suurendab andmete käitlemise tõhusust. Samuti peaks selle tulemusel suurenema metaandmete koosõla laiemalt Euroopas. Metaandmete standard sisaldab statistika metaandmete määratlusi ja esitluse kirjeldusi. See aitab hinnata andmekvaliteeti, dokumenteerida metoodikat ja toota statistikat (EL komisjoni määrus nr 1151/2011).

Riiklikes andmekogudes kasutatakse väljakujunenud klassifikatsioone ja definitsioone, kuid need ei vasta tihti loenduste rahvusvahelistelt harmoniseeritud klassifikatsioonidele ja definitsioonidele, lisaks on loenduse ettevalmistamisel selgunud, et klassifikaatorid, mis on kasutusel riiklikes andmekogudes, ei pruugi olla kirjeldatud RIHA-s (REGREL metoodikaraport, 2013). Oluline töö on koostada loenduse tunnuste metaandmed. Tunnuste kirjelduste kaudu saavad andmekogude pidajad edastada andmeid XML-formaadis, mis võimaldab efektiivset andmete kogumist.

Registripõhine loendus peab vastama loenduse rahvusvahelistele nõuetele, see peab olema (ÜRO, 2006: 7-8):

1. kõikne;
2. ajastatud loendushetkele;
3. kõigi loendatavate suhtes ühetaoline;

4. vaja esitada infot ka väiksemate rühmade ja piirkondade kohta.

2.3 Registripõhise loenduse eelised

Statistikaamet on alates 2009. aastast teinud eeltöid selleks, et korraldada 2020. aasta rahva ja eluruumide loendus registripõhiselt (VVRK, 2009). Registripõhiseks loenduseks on välja kujunenud soodsad tingimused:

- Eestis on registritega kaetud peaaegu kõik Euroopa Liidu kohustuslikud loendustunnused (EL regulatsioon nr 763/2008, vt Lisa 1),
- toimib isikukoodide süsteem, mille alusel on kõik isikuandmete registrid identifitseeritud,
- 2008. aastal loodud ja kasutusele võetud aadressiandmete süsteem on paljudes andmekogudes juurutamisel.

Registripõhise rahva ja eluruumide loenduse korraldamisel on eelised tavalooenduse ees (Tiit, 2015):

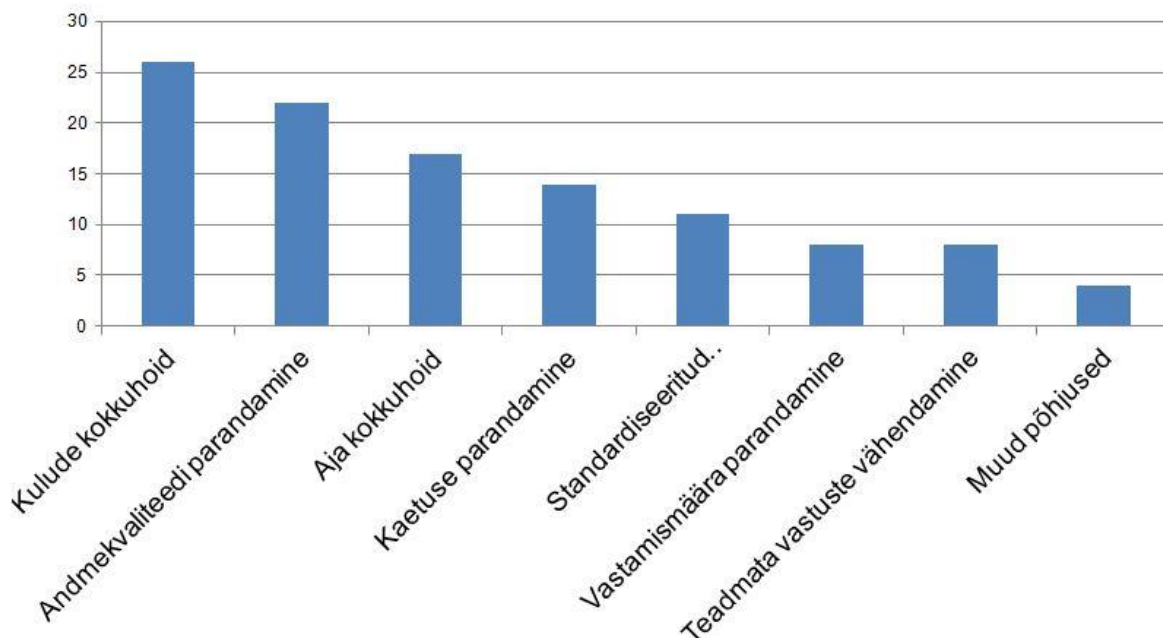
- 1) loendusi saab korraldada tihemini vastavalt vajadusele;
- 2) oluliselt lüheneb andmete korrastamisele kuluv aeg;
- 3) loenduse tarbeks loodavat statistilist registrit hakatakse kasutama teistes statistikatöodes, nt palgastatistikas, rahvastikuarvestuses ja valikuuringutes. Väheneb vastamiskoormus ja vigade tõenäosus, suureneb täpsus ja paraneb kaetus;
- 4) tekib rahvastiku ja majanduse seisukohast oluliste tunnuste aegrida, mis võimaldab protsesse analüüsida ja prognoosida;
- 5) ei ole tarvis andmekogumise välitöid, muutes loenduse odavamaks.

Registripõhist loendust korraldades tuleb arvestada, et sellel meetodil on omad puudused nt andmete *de jure* ja *de facto* vastuolud jne (ÜRO, 2007: 46).

2.4 Registripõhine loendus maailmas

Eesti ja teised Balti riigid kavandavad loendusi järgmiseks loendusvooruks 2020. aastal ilma küsitluseta, üksnes registrite põhjal, nii nagu seda on juba aastakümnete jooksul teinud Taani, Soome, Rootsi, Holland jt (UNECE uuring riikide plaanidest 2020. voo kohta, 2015). Rahvusvaheline kogemus registripõhise loenduse tegemisel ei ole õigusruumide ja registreerimise kultuuri erinevuste tõttu ülekantav. Registripõhise loenduse ettevalmistustööd erinevad nii Eestis, Lätis kui Leedus. Kõik 9 riiki, kus on toimunud registripõhine loendus omavad erinevat laadi andmete registreerimise kohustust, niisamuti on välja töötatud IKT lahendused erinevad. Nt Austrias tähendas registripõhise loenduse ettevalmistus koostöö mudeli loomist kohalike omavalitsustega. Põhjamaades on koosvõime fookuses statistiliste registrite süsteemi vajadused jne (Schulte Nordholt, 2004: 9.22).

Eestil on õppida nendelt riikidelt, sest taotleme registripõhiseks loenduseks loodavalt lahenduselt mitte ainult kuluefektiivsust, aga ka kvaliteeti ning välja töötatud lahenduste mitmekülgset rakendust tulevikus statistika tegemisel laiemalt. Enam levinud põhjused loenduse meetodika vahetamiseks sh registripõhise loenduse meetodi vastu on kulude kokkuvõidust tingitud. Vt Joonisel 1 loenduse meetodika väljavahetamise põhjused riikide arvu järgi (ÜRO, 2011).



Joonis 1. Loenduste meetodika vahetamise põhjused maailmas

Allikas: *UNSD. Report on the Results of a Survey on Census Methods used by Countries in the 2010 Census Round*

Registripõhise loenduskulu väheneb prognoositavalt 10 korda võrreldes tavalise loenduse kuluga (Laihonen,1996). Loendustöö on kallid statistikatöö, seoses loenduse nõudega koguda andmeid kõikide isikute kohta üheaegselt (ÜRO, 2008: 52). Kusjuures loenduskulude kokku arvestamine ei ole lihtne. On välja toodud asjaolu, et loenduskulud riigiti on võrreldamatud, sest puudub ühtne meetodika kuluarvestuseks ja lähtepositsioon loenduse läbiviimiseks on riigiti erinev (nt mõned riigid ei arvesta loenduse ettevalmistava perioodi kulusid loenduskuludesse jne (ÜRO, 2014). Loenduse maksumuse võrdlemiseks maailmas on arvestamiseks kasutatud kahte meetodit eelmises voorus:

1. *Per capita*, rahaühik dollarites;
2. *Per capita*, rahaühik dollarites ostujõu pariteedi alusel 2011. aastal, 31. detsembril.

Eesti kohta on esitatud veebruaris 2013. aastal ÜRO-le järgmised andmed (VVRK, 2013) *Per capita* 13,98 dollarit ja PPP *per capita* 18,56 dollarit.

On teada, et kallimad loendused olid riikides, kus viidi läbi loendus läbi kombineeritud meetodil (Ibid).

Kõige väiksem loenduskulu oli, Sloveenias kus loendus toimus registripõhiselt, kulu *per capita* 0,30 dollarit (Ibid). Kõige odavamad loendused on registripõhised loendused. Loendustööd on Eestis kallimate hulgas maailmas, selle üheks põhjuseks on riigi väike rahvaarv ja mitte piisavad vahendid arendustööde tegemiseks loenduste vahelisel perioodil.

Tabel 1. Kõige kallim loenduskulu loendusmeetodi järgi maailmas 2010. aasta loendusvoorus:

Loendusmeetod	Kokku maksumus (miljon)	PER CAPITA	PER CAPITA (PPP)
traditsiooniline	34.1	5.57	7.31
kombineeritud	18.7	3.94	6.01
registripõhine	1.9	0.24	0.18

Allikas: UNECE 2014. *Measuring population and housing. Practices of UNECE countries in the 2010 round of censuses*

Euroopa kogemus registripõhise loenduse läbiviimisel on positiivne. Parim võimalik meetodika on registrite kasutus loenduseks, sest elanikkonna vastamiskoormus praktiliselt puudub, pole probleeme vastamismääraga, odavam, sest ei sisalda välitöid. (Beltadze, 2015.) Registripõhise loenduse korraldus on riigile odavam võrreldes tavapärase loendusega, kuid meetodika areneb vastavalt registrite süsteemi arengule, kui on tulemas uusi registreid, on avanemas uued võimalused loenduse meetodika arendamiseks.

3. STATISTIKAAMETI ROLL REGISTRIPÕHISE LOENDUSE TEGEMISEL

Loendust korraldab Statistikaamet, kelle ülesanne on pakkuda riigile ja ühiskonnale usaldusväärset ja objektiivset infot (Statistikaameti strateegia 2015-2020). Statistikaameti põhitegevuseks on statistikatööde sh loendusstatistika tegemine. Statistika tegemine peab olema sõltumatu (Euroopa Liidu asutamislepingu artikkel 285). Kõik statistikatööd peavad olema statistika tootmisprotsessis teostatavad ühelaadselt (Euroopa statistika tegevusjuhised). Alljärgnev joonis kirjeldab organisatsiooni põhiprotsessi – statistikatöö tegemist tulenevalt GSBPM mudelist.



Joonis 2. Statistika põhiprotsessi mudel ja selle etapid

Allikas: Statistikaamet

Magistritöös on riigi infosüsteemide koosvõime tagamise teema registripõhiseks loenduseks suurel määral fookuses 4. andmekogumise etapp. Teistel etappidel tehtavad tööd on registripõhisel loendusel ainult statistika tegija vastutusel.

Registripõhise loenduse läbiviimiseks tuleb statistika tegijal ette valmistada infotehnoloogilisel lahendusel põhinev statistika tootmiskeskond ja maksimaalselt automatiseeritud statistika tootmisprotsess, mis võimaldab rahva ja eluruumide loenduse jaoks vajalikud statistikatööd läbi viia registrite baasil (Statistikaameti strateegia 2015-2020 meede 8 ja 12). Registripõhise loenduse ettevalmistustööde fookuses on andmestik ja andmevoog. Andmeid, mida loenduse tulemuste saamiseks kogutakse, on käsitletud tunnustena Euroopa Liidu regulatsioonis ja muudes rahvusvahelistes dokumentides ning statistika seaduses. Tunnustepõhiselt on andmeid kirjeldatud ning nende kvaliteeti hinnatud andmekogudes (REGREL-i metoodikaraport, 2013).

Strateegilise juhtimise koolkonna esindaja Mintzberg on iseloomustanud organisatsioonide strateegiat kui organisatsiooni ja väliskeskkonda vahendavat jõudu, mis sisaldab organisatsioonilisi otsuseid keskkonna suhtes (Mintzberg, 1979: 582-589). Vahendav jõud seoses loendusega seisneb selles, et samm on üldise registripõhise statistika suunas ette võetud, mis nõuab nüüd koostööd andmete omanikega, registripidajatega (Statistikaameti strateegia 2015-2020). Shapiro on toonud välja, et strateegia ei põhine homsetel otsustel, vaid on katse kontrollida tulevikku tänaste otsuste alusel (Katz jt, 1994: 33-115). Statistikaameti strateegia 2015–2020 eesmärgiks on suurendada statistika ajakohasust ja avaldamise kiirust (Statistikaameti strateegia 2015–2020). Kogudes andmeid võimalikult väikese ajakuluga, saab neid avaldada kiiremini, seetõttu on riigil võimalik reageerida omakorda kiiremini ja efektiivsemalt muutustele majanduses.

Statistikaameti strateegiast tulenevad IT-arengukava üldeesmärgid, milleks on loendusega seoses olulised (*Ibid*):

- luua võimalused registripõhise statistika tootmiseks (sh registripõhise rahva ja eluruumide loenduseks), et viia edukalt läbi proovi- ja põhiloendus 2018. a ja 2020. a);
- üle minna küsitlusepõhiselt statistika tootmiselt andmehõivepõhisele statistika tootmisele.

Registripõhise loenduse tuleviku vajadusi silmas pidades on tarvis luua loendusstatistika tegemiseks stabiilne teenus olukorras, kus tekkiva ja kättesaadava info hulk kiiresti suureneb ning oluliseks muutub info, mis tagab andmete kvaliteedi, ja seda saaks kontrollimata usaldada.

Infosüsteemide koosvõime mudel koosneb infosüsteemide teenuste täiendamisest, kasutajaliideste loomises ning suuremahuliste andmehulkade haldamises. Ülesandest tulenevalt on eesmärgid (Hiie, 2014):

- registripidajate rahulolu tagamine;
- tagada stabiilne produktsioon st käideldavus;
- tagada piisav jõudlus st mahuhaldus;
- tagada efektiivne infosüsteemide ülalpidamine (finantshaldus).

Registripõhist loendust plaanides tuleb arvestada, et selle tegemise meetodika on teiselaadne, kuid statistika väljund on ikkagi sama, mis tavalisel viisil läbi viidud loenduste korral (ÜRO, 2007: 47). Registripõhise loenduse puhul peab statistika tegija arvestama asjaoluga, et registrid on loodud administratiivsetel eesmärkidel ja statistika tegemiseks on vaja mitmepoolset koostööd ja pakkuda tuge andmekogudele nt kvaliteediprobleemide lahendamiseks, niisamuti registripõhise loenduse meetodika väljatöötamiseks.

4. PROBLEEMI ANALÜÜS

4.1 Regstripõhise loenduse eellugu

Regstripõhise loenduse korraldamisega seonduvat probleemi uurisin dokumenteeritud andmeallikatest:

1. Statistikaameti ametkondlikku kirjavahetust ministeeriumitega registrite teemal loenduse kontekstis perioodil 2000–2009;
2. Statistikaameti koostatud analüüsid perioodil 2007-2011 registrite kvaliteedi kohta;
3. Paber kandjal analüütiline ülevaade „Registrid ja klassifikaatorid rahvaloenduses. Automaatkodeerimine” 1996. aastast, mille koostas Jüri Laast Ph. D., AS Infoekspert;
4. REGREL projekti metoodika raport 2013. aastast;
5. Vabariigi Valitsuse rahvaloenduskomisjoni protokollid perioodil 2008–2013;
6. Vabariigi Valitsuse loenduskomisjoni protokollid 2015. aastal;
7. REGREL teemalised ettekanded perioodil 2014-2016 aastal.

Väga mitmekülgne ja ajastukriitiline teemakäsitlus registriandmetest on esitatud perioodi 2010–2013 regstripõhise loenduse vajadusest lähtuvalt „REGREL-i metoodikaprojekti” raportis, mille koostajateks on Eesti teadlased. Nimetatud raport pakub juhiseid regstripõhiseks loenduseks ettevalmistamisel õiguse, semantika ja andmekvaliteedi parandamise teemadel. Antud töö keskne väärtus on põhjalik registrite andmete analüüs rahva ja eluruumide loenduse tarbeks.

Allikate analüüsi lähteülesandeks oli välja selgitada:

1. Millal esimest korda on mainitud registrite kasutamise kavatsust loenduseks?
2. Millised on olnud aktuaalsed infosüsteemide koosvõimekuse teemad?
3. Milline on olnud riigi infosüsteemide ja registrite andmekvaliteedi areng ajas?

Analüüsi ülevaate esitan kahe perioodi kohta, murdepunktiks on subjektiivselt määratud regstripõhise loenduse projektorganisatsiooni loomine 2009. aastal:

Esimene periood: 1996 kuni 2009 aasta.

Teine periood: 2010 kuni 2015 aasta.

Registripõhise loenduse diskussioon on olnud päevakajaline seoses vajadusega alustada loenduste ettevalmistustöödega. See on seotud olnud ressursside planeerimisega loendustöödeks. Loenduste ettevalmistamise käigus on püütud eraldatud eelarveliste vahenditega arendada jätkusuutlikkuse põhimõtetel infosüsteeme. Kuid IT-tehnoloogia areng on olnud kiirem perioodil 1995–2011 ja platvormid, süsteemid, mis võeti kasutusele 1998, ei olnud efektiivselt kasutatavad viimasel loendusel 2011. aastal. See ongi olnud määrav asjaolu, miks loenduskulud on perioodil 2000-2011 1,6 korda kasvanud. Kuluartiklitest on suurimad kulud 2000. aasta ja 2011. aasta loendusel seotud andmekogumise läbiviimisega (loendajate palgad) ja statistika tootmissüsteemi (infosüsteemide loomisega) ettevalmistamisega (VVRK, 2008).

Registripõhise loenduse korraldamise võimalikkust on vaagitud alates 1996. aastast, kui loenduse tegemisi juhtis L. Tepp. Ta andis resoluutselt 90ndate keskpaigas teada, et Eestis puuduvad usaldusväärsed registrid, sest neid alles luuakse, ning registrites olevad andmed on lünklikud ja lisaks pole andmeid tegelikkuses kontrollitud (Laast–Laas, 1996). Kokkuvõtlikult tõdes L. Tepp, et registreid pole võimalik loendusel kasutada ning 1998. aastal alustati traditsioonilise loenduse ettevalmistustöödega riigis. Sellest ajastust on pärit dokument, kus on kaardistatud erinevatest loenduse meetodikatest tulenevalt IT-tehnilisi vajadusi. See on välja töötatud AS Infoekspert ekspertide poolt. Dokumendi välja töötamise kavatsuse kohta puudub taustinfo, võib oletada, et see oli seotud IT-tehnilise baasi, rahvaloenduse infosüsteemi, loomisega Statistikaametile. Nimetatud dokumendis on esmakordselt mainitud võimalust registrite kasutamiseks loenduseks ning esitatud kava registrite info puuduste korvamiseks (Laast–Laas, 1996). Kahjuks ei leidnud kava rakendust, võimalik põhjus, et tähelepanu fookuses olid loenduse tööd ja lisategevusteks polnud ressursi. J. Laast–Laas oli seisukohal, erinevate registrite ühildamine on suuremate või vähemate jõupingutustega võimalik (Laast–Laas, 1996: 81).

Ametlikku kirjavahetust registrite teemal seoses loendusega ega ka andmemetoodiliste tööde analüüse ei leidu loenduse 2000. aasta tegemise ja lõpetamise ajal 2004. aastal. Registrite teema tõstatub ametkondade kirjavahetuses taaskord 2005. aastal, sellest aastast võib

täheldada süstemaatilist tegevust registrite kvaliteedi hindamisel. Miks registrite teemat ei käsitletud võib seostada sh range kehtiva õigusruumiga statistika tegemiseks, mille järgi nt 2000. aasta loendusandmete kasutamine analüüsiks oli seaduse alusel keelatud (kehtis kuni uue seaduse jõustumiseni 2010. aastal) selle tõttu polnud võimalik ka andmete valideerimise tööd registriandmetega. Nt Statistikaameti tehtud 2003. aasta rahvastikuregistri kvaliteedianalüüsi kasutamine oli keelatud Andmekaitse Inspektsiooni poolt (Tikk, 2008). 2005. aastal Statistikaametil õnnestus selgitada inspektsioonile vajadust seoses rahvastikustatistika tegemisega ning uurida rahvastikuregistri tunnuste täituvust, andmete kvaliteeti ning võrrelda definitsioone (ETU ja rahvastikuregistri elukoha analüüs 2005. aastal). Statistikaameti seisukoht registri andmestiku kvaliteedi kohta esitati 2006. aastal Rahandusministeeriumile ning toodi välja probleemid, mis seisnesid selles, et registri eesmärk ei ole anda andmeid statistika tootmiseks (andmekogupidaja kogus statistilisi tunnuseid vabatahtlikkuse printsiibile vastavalt, registritoimingud baseeruvad teisel ärioloogikal ning lisati olemasoleva info kvaliteedi kohta kriitiline hinnang, mis seostati asjaoluga, et registripidaja ei rakendata süstemaatiliselt loogilisuse kontrollle ning seetõttu on olemasolev info registris sageli puudulik (Tiit, 2009). Samal aastal edastas Statistikaamet Majandus- ja Kommunikatsiooniministeeriumile kirjalikud ettepanekud tegevuste kohta registrite kvaliteedi parandamiseks seoses kavandatava loendusega 2011. aastal ning perspektiiviga registripõhise loenduse korraldamiseks tulevikus (Ehitisregistri kohta küsimused ja vastused 27.03.2006).

2007. aastast on fikseerida uus algatus. Statistikaamet analüüsis rahva ja eluruumide loenduse näitajate olemasolu riiklikes andmekogudes registrite enesehindamise küsimustiku põhjal. Uurimusest järeldus registriandmete kasutamise otstarbetus tulenevalt koostatud loendusprogrammist, seda põhjusel, et andmekogude ja registrite täitumus oli loenduse jaoks vajalike näitajate osas mitte piisav (Paut analüüs, 2008). Täielikult puudusid leibkondade koosseisu ja tegeliku perekonnaseisu, tööaja, usu, võõrkeelte oskuse andmed ning andmed leibkondade põllumajandusliku tegevuse kohta. Kogu elanikkonna kohta ei olnud andmeid ameti, emakeele, leibkondade ja leibkonnaliikmete vaheliste perekondlike seoste, naistel sünnitatud laste arvu, omandatud haridustaseme, rände, vanemate sünnikoha kohta. Puudulikud olid elamistingimusi iseloomustavate näitajate andmed. Teadmata oli isikute tegeliku elukoha ja registreeritud elukoha ühtivus. Puudus ülevaade institutsioonides (kloostrid, lastekodud jt) elavate inimeste kohta. Ei olnud võimalik siduda rahvastikuregistri

ning ehitiste ja hooneteregistrit, lisaks võimalik oli olukord, kus inimene on registreeritud eluruumi, mida ei eksisteerinud pealegi olid metaandmed registrites puudulikud. (Tiit analüüs, 2009.)

EL tõukefondide avanemisega on Eesti andmekogude jaoks ära märkida positiivne edasimineku 2007. aastal. Euroopa Liidu Inspire direktiivist tulenevate kohustustega seostada ruumiandmed koha-aadressidega kutsuti ellu Maa-ameti projekt ERAS. Projekti eesmärk oli seostada Eesti Topograafilise Andmekogu ehitiste ruumiandmed koha-aadressidega ning lisada hoonetele rahvaloenduseks vajalikud atribuutandmed nt eluhoone, tööpaik, institutsioon. Probleemseks kujunes selle töö juures 2008. aastal reaalses ruumis teadmata aadresside asukoht, seda põhjustas asjaolu, et katastriüksuse ja Eesti topograafilise andmekogu hoonete andmed ei olnud seotud. (Statistikaameti memo ümarlaua kohtumisest aadressiandmete teemal 2008. a.) 2008. aastal algatas Maa-amet uue projekt AKP, mille tulemusel pidi riik saama korrastatud koha-aadressid ja aadressiobjektide ruumiandmed. Vastuvõetud andmekorrastustöö kvaliteet osutus reaalsuses kontrollimise käigus mittepiisavaks (VVRK, 2008). Aadressiandmete korrastustöödega oli vaja jätkata hiljem peale loendust 2011. aastal.

Registripõhise loenduse koosvõime temaatika muutus taas aktuaalseks eelmise loenduse (2011. aasta) meetodi valiku kinnitamisel, taaskord seoses loenduse kulude prognoosiga. Loenduskulude maksumuseks kinnitati 340 miljonit krooni (VVLK, 2008), seda asjaolu arvestades sai detsembris 2008. aastal Statistikaamet ülesande Vabariigi Valitsuse rahvaloenduskomisjonilt koostada andmekogude korrastamise koordineerimiseks lähteülesanne rahva ja eluruumide loenduse miinimumvajaduste (Euroopa Liidu nõuded) kontekstist lähtudes (VVRK, 2008). Statistikaameti esitatud lähteülesande alusel tuli Majandus- ja Kommunikatsiooniministeeriumil koostada koostöös rahvaloenduseks olulisemate andmekogude omanikega ja Statistikaameti esindajatega analüüs, mis pidi kaardistama ja välja selgitama õigusliku raamistiku ning registreeritud andmekvaliteedi parandamise võimalused (VVRK, 2009). Statistikaamet sai lisaülesande Vabariigi Valitsuse rahvaloenduskomisjonilt 2009. aastal, mille sisuks oli kohustus kasutada registriandmeid rohkemates loendustöödes kui ainult üldkogumi moodustamine. Statistikaamet otsustas registriandmeid kasutada loenduse ankeedi eeltäitmiseks (*Ibid*).

Perioodi 1996-2009 kokkuvõtteks on märkida, et eksisteeris kaks suunda registrite kasutamise teemadel loenduseks: pessimistlik, mis toonitas andmekvaliteedi puudulikkust ja suuri andmelünkasid registrites ning optimistlik suund, mis tõi välja võimaluse loenduse kulusid kokku hoida, kui kasutada selleks kaasaegseid infotehnoloogilisi vahendeid, erinevate registrite andmestiku infointegreerimise võimalusi, erinevate asutuste hajutatud infotööd ümber korraldades. Pessimistliku suuna esindajate arusaam lähtus vastutusest loendustulemuste kvaliteedi eest, lünkadest seadusandluses, riist-ja tarkvaraplatvormide sh andmebaaside ja süsteemide puudulikkust võimalusest neid ajakohastada statistika tegemiseks. Samas avaldati soovi ja valmidust teha rohkem koostööd erinevate asutuste ja ametkondadega, et uurida registrite andmeid ja leida võimalusi olukorra parandamiseks. Registrite andmekvaliteedi parandamiseks soovitas statistika tegija fikseerida kvaliteedinõuded, viia sisse registrite auditeerimine. Ühtlasi tuletati meelde vajadust, et registrites tuleb tagada andmete regulaarne aktualiseerimine ning hädavajalik on ühtlustada erinevates registrites kasutatavad definitsioonid ja klassifikaatorid.

Ülevaade registrite info koosvõimet puudutavates küsimustes on esitatud alljärgnevas tabelis 2.

Tabel 2. Kokkuvõtte registripõhise loenduse teemapüstitusest perioodil 1996-2009

Probleemid	Probleemide põhjustajad	Nõuded	Tegevussuunised
Rahvastiku ja ehitiste register ei kasuta ühtset aadressi esitust.	Registreid oli arendatud eraldi.	Erinevad registrid peavad ühilduma identifikaatorite osas	Töötada välja aadresside standard ja võtta see kasutusele kõigis riiklikes registrites; sealjuures fikseerides vajadusel isikule mitu aadressi.
RR hariduse skaala (klassifikaator) erineb oluliselt kasutatavast, hariduse infosüsteemi klassifikaatorist.	Eesti seadusandluses on alatähtsustatud registripidamist, kirjete kandmine registrisse/ registri täiendamine või parandamine on jäänud mõne (vastutust mitte kandva) isiku suva hooleks (isikute registreerimine eluruumi, hoonete registreerimine).	Kõik registrid kasutavad ühesuguseid definitsioone ja klassifikaatoreid	Legaliseerida 2000. a loenduse andmebaasi kasutamine registrite kontrollimiseks ja täiustamiseks.
Rahvastikuregistri täieliku objektide	Andmekaitse on oluliselt taksitanud	Registrid peavad olema katvad	Kohustada kõiki registreid hindama

loetelu kohta võib oletada ülekaetust.	registrite ristikasutamist, mis on peamiseks teeks registrite parandamisel. Pole olnud võimalik kasutada 2000 loenduse andmebaasi andmete valideerimiseks.	uuritava populatsiooni osas ja neid ajakohastatakse pidevalt.	nende kvaliteeti ühtsetel alustel (kvaliteedistandard) ja töötama kvaliteedi parandamise nimel.
Registri andmete kontrolliga tegeleb peamiselt Statistikaamet.	Puudub huvi registriandmete kvaliteedi hindamiseks, puuduvad kvaliteedistandardid, mille alusel saaks registrite andmekvaliteeti hinnata	Registrite kvaliteeti kontrollitakse pidevalt ajas ja registripidajate ülesanne on vigade leidmine, avastamine ja parandamine.	Ajakohastada loendusseadust 2000 aastast.

4.2 Registri põhise loenduse ettevalmistamine

Uue hoo sai registripõhise loenduse temaatika seoses Riikliku statistika seaduse rakendamisega 2010. aastal. Statistikaamet sai kohustuse alustada registripõhise loenduse ettevalmistustöid. Seaduse tasandil fikseeriti ülesanne võrrelda 2011. aasta loendusandmeid vastu registriandmeid (Riiklik statistika seadus, loenduse peatükk). Andmekogude ja registrite täituvus jäi enne eelmist loendust vajalike näitajate osas puudulikuks (VVRK, 2011). Lisaks (Tiit analüüs, 2011), vt Lisa 3:

1. Rahvastikuregistri ja ehitisregistri aadressid ei ühildunud;

2. Registritel polnud andmekvaliteedi kohta ülevaadet;
3. Põhiregistrites (RR, EHR) ei rakendatud süstemaatiliselt kontrolle (nt loogilisus), mistõttu info oli sageli puudulik või oli vale.

Loenduse andmetega registrite kvaliteedi hindamine ja parandamine sai eelduseks edaspidisele täielikult registripõhisele loendusele üleminekuks. Andmestike võrdlusanalüüs tehti ära 2014. aastal, peale seda kui kõik loendustulemused olid avaldatud. Analüüsi töö tulemusel tehti järeldus, et kui kombineerida registrite ja loendusandmeid saavutatakse efekt, mille tulemusel registritest pärit andmete andmekvaliteedi näitaja mitmete tunnuste puhul paraneb loenduse andmete töötlemise protsessis. Selle järelduse kohaselt registreid saab kasutada loenduse tegemiseks (Tiit, 2015). Samas analüüsitulemused sarnaselt meetodikaraporti tulemustele 2013. aastast viitasid andmete kvaliteedi ebapiisavusele. Vt tabel 3. Isiku – ja leibkonna ning eluruumi tunnused registrites ja Lisa 6, kus toodud välja kvaliteedi probleemid loendustunnuste kategooriates.

Tabel 3. Isiku- ja leibkonna ning eluruumi tunnused registrites

Hinnang kaetusele registrites	Loendustunnused
A – täielik	Sugu, vanus, sünniriik, kodakondsus
B – osaline	Hõiveseisund, tööalane seisund Kasulik pind/tubade arv, eluruumi hoone tüüp
C – piiratud	Alaline elukoht, välismaal elamine ja riiki saabumise aasta, eelmine elukoht, leibkonnaliikmete vahelised seosed, leibkondade valduse staatus, majanduse tegevusala (Parandus läbi õigusruumi 2015–2016)* Eluaseme korraldus, eluruumi tüüp, eluaseme korraldus, tavaeluruumide valduslik olukord, eluruumide kasutamise alus, omandisuhe, elanike arv eluruumis, eluruumi tehnilised tunnused (veevarustussüsteem, tualettruum, pesemisvõimalus, kütmistüüp, ehituse aeg) (Ehituseaduse muudatus alates 1.07.2015) *

D - puudub	Amet, töökoha asukoht (Vajadus läbivaatamisel riigi projekti „Aruandlus 3.0 raames“)*
------------	---

*Autori märkus

Allikas: REGREL metoodikaraport 2013. aastal

2013. aasta metoodikaraportis on viited vajadusele alustada töid andmekvaliteedi parandamiseks registrites ent vastusega jäi küsimus, kuidas kogutavaid andmeid registrites peaks saama parandada. Arvestades, et registrid on Eestis dokumendipõhised, mis seab piirangu korrastustöödeks, sest vaja on õiguslikku alust. Õigusloome selleks on keeruline, sest tegelikult ei saa olukorda lahendada ei registripidaja ega ka Statistikaamet, häid lahendusi hetkel ei ole. Teise perioodi 2014. aastaks oli välja kujunenud uus tase riigi infosüsteemide koosvõimekuse teemal:

- 1) registritega oli kaetud praktiliselt kõik loendustunnused ja amet ja töökoha asukoht;
- 2) toimis isikukoodide süsteem, mille alusel enamus isikuandmete registreid oli identifitseeritud;
- 3) oli olemas aadressiobjektide koodid ja kasutusel äriregistri kood.

Tulenevalt uuest olukorrast, otsustas Statistikaamet ette valmistada registripõhine loendus ning testida olukorda kahel prooviloendusel 2016. ja 2018. aastal (VVLK, 2015).

Püstitati uued lahendamist vajavad ülesanded (Tiit, 2015):

- Lahendada sisuline/ juriidiline/ tehniline registripõhise loenduse võimalikkus.
- Lahendada registri- ja loendustunnuste definitsioonide vahekord.
- Luua loendusandmete hõivekontseptsioon ja plaan.
- Luua metoodika, mille alusel määrata loenduse üldkogum st keda loendada.
- Luua kvaliteedi parandamise raamistik, mis tagaks kaetuse, täpsuse, andmete regulaarne värskendamise registrites.

Koostoimeks erinevate andmekogude pidajatega on koostatud tegevuskava, mille on heaks kiitnud Vabariigi Valitsuse loenduskomisjon (VVLK, 2015). 2015. aastast on registripõhise

loenduse diskussioon suunatud infosüsteemide koosvõimele, andmestiku kvaliteedile (VVLK, 2015). Positiivsena võib märkida, et alates 2016. aastast on eelarvelisi vahendeid rahvastikuregistri andmekvaliteedi parandamiseks ja augustist, 2015 aasta, on avatud eraldi tõukefondide voor REGREL arendustööde rahastamiseks Statistikaametile ja andmekogude pidajatele (Majandus ja taristuministri käskkiri nr 15-0213, 14.06.2015).

Kokkuvõtvalt võib öelda registripõhise loenduse diskussiooni uurimise perioodil 2010-2015 on toimunud areng infosüsteemide võrgustamise ja koosvõime suunal. Aadressistandard on kasutusele võetud rahvastikuregistris, andmekvaliteedi mõõtmisesse panustavad nii andmekogupidajad kui ka Statistikaamet, on olemas turvaline andmevahetuskeskkond, registritel on omanikud ja nendel ülesanded, loendusandmeid võib kasutada andmemetoodilisteks töödeks, on kehtestatud kvaliteedi standardid andmetele.

Kahe perioodi võrdluses on probleemikohad registripõhise loenduse korraldamisel läbi aastate 1996-2015 olnud sarnased. Nendeks on:

- 1) loendusandmete kvaliteet;
- 2) puuduvad andmed;
- 3) registrite ühilduvuse infotehnoloogiline tagamine (identifikaatorid, klassifikaatorid);
- 4) aadressiandmed;
- 5) õiguslik raamistik.

Vähesel määral leidis registripõhise loenduse diskussiooni raames käsitlust infosüsteemide andmeturbe teema ja andmevahetus üle x-tee.

Alates 2015. aastast on aktuaalsem teema registripõhise loenduse tegemisel andmete sisu ja kvaliteet (VVLK, 2015):

1. Registreeritud ja tegeliku elukoha andmete erinevus registrites ligi veerandil elanikkonnast.
2. Ametiala ja töökoha asukoht puuduvad registrites kõigi töötajate kohta.
3. Ehitisregistri andmete kvaliteet pole piisav statistika tegemiseks. Ehitisregistri madal kvaliteet sisaldab nii hoonete kui eluruumide alakaetust kui ka tehniliste tunnuste puudulikkust täidetust ja aktualiseerimatust.

Registripõhise loenduse võimalikkust on vaagitud juba 1996. aastast, üldistatult võib öelda, et probleem on sama vana kui Eesti riigi registrid. Aktuaalsemad teemad on olnud seoses standardite (aadressiandmete, klassifikaatorite) kasutamisega seoses niisamuti loenduse jaoks puuduvad tunnused registrites. On rõhutatud, et registripõhise loenduse meetodi valik eeldab, et registritega on kaetud praktiliselt kõik kohustuslikud loendustunnused; toimib süsteem, mille alusel on kõik loenduse vaatluse objektid identifitseeritud; kasutatakse aadressandmeid.

4.3 Registripõhise loenduse õiguslik raamistik

Registripõhise rahva ja eluruumide loenduse korraldamiseks vajaliku õigusruumi hindamiseks on tehtud mitu õiguslikku analüüsi. Esimene 2008. aastal riikliku statistika seaduse uue redaktsiooni välja töötamise valguses (Tikk, 2008).

Teine suurem töö toimus 2012. aastal REGREL projekti käivitamisel kui analüüsiti andmekogude õigusruumi, et selgitada välja Euroopa õiguses sätestatud kohustuslike loenduse tunnuste moodustamise võimalikkus andmekogudes olevate andmete alusel ja kaardistati võimalikud kitsaskohad andmete kogumisel andmekogusest ning õiguslike muudatuste ulatus (REGREL metoodikaraport, 2013). Kolmandat korda analüüsiti 2014. aastal andmekogude põhimäärusi ja nende unifitseerimise võimalusi lähtuvalt kehtivast õigusest (Künnapas ja Lindberg, 2014).

Esitan uurimisküsimuse: kas õiguslik raamistik registripõhise loenduse korraldamiseks on olemas? Küsimus on seotud ankeetküsitlusest ilmnenu vastustega, mille järgi puudub andmekogude koosvõimeks õiguslik alus. Vt Lisa 10.

REGREL metoodikaprojekti raames viidi 2010–2013 läbi kaheosaline õigusanalüüs. Projekti esimeses osas tehti õigusruumi kontentanalüüs, mis vaatles õiguslikke võimalusi andmete töötlemiseks statistika vajadusteks lähtuvalt statistika valdkonna õigusaktidest. Selle analüüsi osa keskendus riikliku statistika seaduse, isikuandmete kaitse seaduse ja avaliku teabe seaduse eritlemisele (REGREL metoodikaraport, 2013). Analüüsis vaadeldi registripõhise loenduse võimalike lahenduskäikude õiguslikke mõjusid. Metoodikaprojekti REGREL lõpparuande

kohaselt on registritepõhise loendusega ja laiemalt andmekogudesse kogutud andmete statistika vajadusteks kasutamiseks seotud kolm spetsiifilisemat probleemi (*Ibid*):

1. Rahva- ja eluruumide loenduseks vajalike ning registrites olevate andmete erinev tähendus nt sünnikoht, tegelik perekonnaseis. Probleemi lahendamine eeldab tööd definitsioonide harmoniseerimiseks.
2. Registrites olevate andmete kvaliteet (sh kaetus ja aktualiseeritus). Probleemi lahendamine eeldab andmekorrastustöid andmekogude pidajatelt.
3. Registrites puuduvad andmed, kuna need andmed ei ole olnud vajalikud ühegi konkreetse haldusmenetlust läbi viiva haldusorgani ülesande täitmiseks (Eestis ametiala ja töökoha asukoht). Probleemi lahendamine eeldab kokkuleppeid andmesaadajatele kohustuse tekitamiseks.

2014–2015. aastal viidi läbi REGREL õigusanalüüs, mille eesmärgiks oli analüüsida andmekogude pidamise aluseks olevat õigust ja selle tulemusel välja töötada õigusaktide muudatusettepanekud, mis tagaksid registripõhise loenduse läbiviimise. Analüüsi tegijad tõdesid, et õiguslikest aspektidest probleeme ei ole registripõhise loenduse tegemiseks, kuid vajalik on saavutada ametkondlikud kokkulepped andmekoosseisude kohta. Õigusanalüütikud on välja selgitanud, et Statistikaametile on tagatud õigus kasutada andmete töötlemisel isikut identifitseerivaid tunnuseid ning riigi andmekogude andmeid (Tikk, 2008). Lisaks Statistikaametil on kohustus kasutada statistika tegemisel eelkõige riigi infosüsteemidest andmeid. Sellega seoses on uuritud unikaalsete andmete kasutusest. AvTS-s järgi on riigi infosüsteemi kuuluvasse andmekogusse kogutavad andmekogu unikaalsed andmed põhilandmed, mis tekivad andmekogu haldaja avalike ülesannete täitmise käigus (Avaliku teabe seadus, andmekogude peatükk). Riigi infosüsteemis sisalduvate andmete kvaliteet peaks andmeid põhilandmetena töötlevates andmekogudes olema tasemel, mis võimaldab teistel andmesaadajatel andmekogu põhjal neile pandud ülesandeid täita.

Toetudes ankeetküsitluses välja toodud kõige olulisemale takistusele andmekogude pidajate poolt andmekvaliteedi tagamine andmekogudes registripõhise loenduse tegemisel, tuleb siinkohal nentida õigusliku raami olemasolu ei ole taganud andmekogude töötlejate reaalset suutlikkust AvTS-is sätestatud kohustusi täita andmekvaliteedi tagamise ulatuses (Lisa 10). REGREL-ile üleminekul lähtutakse riikliku statistika seaduse RSS §-s 50 sätestatust (Riiklik statistika seadus loenduse peatükk). Selle kohaselt hindab Statistikaamet andmekogude

põhiandmete kvaliteeti ning teeb vajaduse korral andmekogu vastutavale töötlejale ettepanekuid andmete kogumiseks ja/või kvaliteedi parandamiseks. Lahendamatuid õiguslikke probleeme selle teema osas pole, kui mitte arvestada asjaolu, et õigusaktidest selgub üheselt, et riikliku statistika tegija ülesanne on järelevalve teostamine andmete kvaliteedi kohta andmekogudes seoses statistika tegemisega, missugune on selle ülesande ulatus on eraldiseisev probleem. Selgusetuks on jäänud, kes koostab kvaliteedi standardid andmekogudele riigis. Välismaa registripõhise loenduse korraldust silmas pidades, tuleb nentida, et ainult üksikud riigid riskasutavad registrite andmeid. ESS süsteemis on detailset fikseeritud nõuded, mis seotud isikustatud andmete kasutamisega (Euroopa statistika tegevusjuhised). Kerkib esile küsimus, milline on/või peaks olema tagasiside formaat andmekvaliteedi kohta andmekogule. Selle kohta pole eksisteerivat normi.

Õigusanalüüsid viitavad õigusraamistik on registripõhiseks loenduseks olemas, tõlgendamise ruum tuleb läbi rääkida ja tavapärase haldusorganite koostöös lahendada (Künnapas ja Lindberg, 2014). Kuid andmekvaliteedi parandamiseks on riiklikel andmekogude pidajatel töö, mida pole võimalik teha selgepiirilise õigusnormita. Näiteks, kui hoone omanik on eelmise sajandi alguses juriidilise dokumendi alusel fikseerinud registris andmed ning tänaseks on ta mõne teise riigi resident, siis pole omanikul kohustust uurida ja täpsustada parandada temale kuuluva omandi andmeid ehitusregistris. Nii eksisteerib andmebaasis andmete hulk, mis ei vasta tänapäeva andmekvaliteedi nõuetele, mis omakorda tekitab probleeme statistika tootmiseks, kuid seevastu on andmed õiguslikult kehtivad.

Riigis on kehtestatud raamistike näol soov andmekogude omavahelist koosvõimet suurendada, mis eelkõige tähendab andmevahetust (andmete riskasutust) andmekogude vahel, samas on protsessid ühepoolset õiguslikult reguleeritud (Riigi infosüsteemide koosvõime raamistik, 2011). Teema selgitamiseks tuleb korraga teha ajalooline kõrvale põige. Varasem andmekogude seadus sätestas andmete riskasutuse mõiste, milleks on andmete ülekandmine ühest andmekogust teise või mitmes andmekogus sisalduvate andmete ühine infotehnoloogiline töötlemine. Kehtivas õiguslikus regulatsioonis on mõistest loobutud, toimub andmevahetus andmekogude vahel ning selles kontekstis väärib märkimist, et kui ühe riigiasutuse sisesteks vajadusteks loodud andmekogu on arenenud niisuguseks infosüsteemiks, mis võimaldab aidata kaasa ka teise riigiasutuse ülesannete tõhusamale täitmisele, siis peab andmekogu regulatsioon tervikuna toetama andmete edastamise õigust (Avaliku teabe seadus, andmekogude peatükk). Seega on riskasutus isikuandmete kaitse

õiguse seisukohalt tänaseks ajalooline mõiste, mis kaotas oma aktuaalsuse uue andmekogude regulatsiooni vastuvõtmisega. Tehniliselt on riskasutus üks isikuandmete töötlemisjuhtum, mis on kaetud töötlemise üldmõistega isikukaitse seaduses. Sisuliselt aga loovad riskasutuse õigusliku aluse sätted, mis annavad andmekogudele õiguse oma ülesannete täitmiseks teistelt isikutelt, sh andmekogude töötajatelt, andmeid saada (Tikk, 2010). Võib tuua välja, et vaatamata asjaolule, et mitmed riigi andmekogud (maksukohustuslaste register, liiklusregister jt) on rahvastikuregistriga riskasutuses ja nendes lähtutakse füüsiliste isikute puhul rahvastikuregistri andmetest, koguvad mitmed andmekogud rahvastikuregistri põhiandmeid nt elukohaandmed dubleerivalt enda poolt peetavatesse registritesse. Selline käsitlus registripõhise loenduse ettevalmistamiseks ei ole koosvõime tagamise aspektist hea, kuid see on tingitud asjaolust, et pole piisavat tähelepanu andmekogumise dubleerimise järelvalvele. Oleks vaja rakendusmeetmeid, et riigis oleks efektiivsemalt korraldatud kogutavate andmete jagunemine erinevate andmekogude vahel, juba kogutud andmete vastavuskontroll, riskasutus ning vastuolude kõrvaldamise protsess kogutud andmete vahel. Õiguslikku alust järgides võib andmekogudesse koguda ka dubleerivaid andmeid, kuna Avaliku teabe seaduse järgi võib andmekogus töödeldavate korrastatud andmete kogum koosneda ka üksnes teistes andmekogudes sisalduvatest unikaalsetest andmetest. See on eesmärgipärane, kui haldusorganil on oma ülesannete täitmiseks vaja andmeid, mida ta peab kasutama, siis menetlustoimingute tegemiseks, tõendite esitamiseks jne on see tegevus asjakohane. Keelatud on asutada ühtede ja samade andmete kogumise eesmärgil eraldi andmekogusid (Avaliku teabe seadus, andmekogude peatükk). 2020. aastal läbiviidava rahva ja eluruumide loenduse eripäraks on selle läbiviimine andmekogudesse kogutud andmete põhjal, et tagada halduskulude kokkuhoid ning andmesubjektide minimaalne koormamine andmete esitamise kohustusega. Registripõhise loenduse õigusliku aluse annab RStS § 28 lg 4 koostoimes RStS § 29 lg-ga 18 (Künnapas ja Lindeberg, 2014).

REGREL-ile üleminekul lähtub Statistikaamet riikliku statistika seaduse §-s 50 sätestatust. Väärrib mainimist, õiguslase dokumentatsiooni analüüsi käigus ilmnes asjaolu. Nimetatud §-s 50 sätestatu järgi ei ole reaalne registripõhist loendust korraldada, sest paljusid tunnuseid nt murre, religioon jt pole olemas registrites, neid tuleks eelkõige hakata koguma registritesse. Peale selle on loenduse korraldust käsitlev paragrahv üles ehitatud põhimõtetal, et loendused toimuvad traditsioonilisel viisil, kus andmeesitajaks on füüsiline isik. Registri põhisel loendusel on andmeesitaja registripidaja. Antud informatsiooni valguses on ettepanek Riikliku

statistika seaduse ajakohastamiseks loendust käsitleva paragrahvi ulatuses. Autor on arvamusel, et andmekogude koostoime osa on tagatud raamistikega, kuid see pole piisav (Avaliku teabe seadus §43). Statistika tegijal pole tagatist, et andmekogude töötledjad suudavad tagada loenduse läbiviimiseks vajaliku kvaliteediga andmete olemasolu ja edastamise.

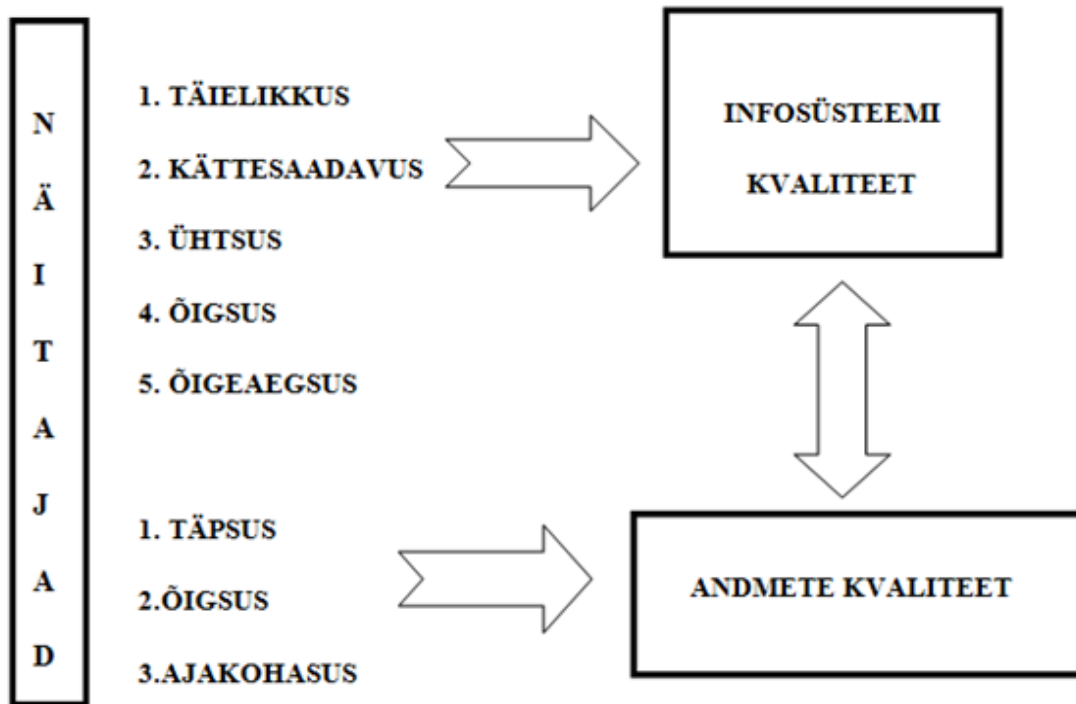
5. RIIGINFOSÜSTEEMID REGISTRIPÕHISEKS LOENDUSEKS

Antud peatükis keskendun võtmemõistete komplektile, mis seotud riigi infosüsteemide ja loendusega. Peatükis on kesksel kohal mõisted: infointegratsioon, infosüsteem, koosvõime. Samuti toon välja infosüsteemide erinevatest eesmärkidest ja arengutasemetest tulenevad probleemid.

Süsteemi teoreetiliselt on defineeritud infosüsteemi kui omavahel seostatud ja mitmel tasemel eksisteerivate andmete, funktsioonide, protsesside, sündmuste, asukohtade ja väärtuste süsteem

(Mikli, 1998). Käsitlen peatükis mõistet infosüsteem kui vahendit, mille struktuur põhineb eeskirjadel ning toimimine on suunatud info hulga kogumise, säilitamise ja väljastamise loogikale (Mereste, 1987). Inimene on sh infosüsteemi osa ja koostoimeks on vajalik tagada organisatsiooni struktuuri loogiline ühilduvus infosüsteemidega ja seotus tootmisprotsessiga (Hiie, 2014). Kui muuta infosüsteemis ükskõik millist komponenti võib muutuda ka organisatsioon, kus inimesed kasutavad infosüsteemi sestap tuleb hinnata infosüsteemi kvaliteedi näitajaid täielikkuse, kättesaadavuse, pidevuse, ühtsuse, õigsuse ja õigeaegsuse kohta (Alberts jt, 2001). Lisaks kvaliteeti hinnangule on loenduse statistika tegemise aspektist väga oluline andmekvaliteedi tagamine nt täpsus, õigsus, ajakohasuse näitajate kohta (ÜRO, 2006:17-90). Madal andmekvaliteet võib mõjutada erinevate infosüsteemide koosvõimet, kuid samas ei pruugi olla infosüsteemide ühilduvus probleemide tekitajaks (Viilup, 2011). Näiteks, ajakohastamata metaandmed loendusstatistikas võivad põhjustada tõrgetega andmetöötluse, kuid andmeedastus ühest infosüsteemist teise kulgeb sujuvalt. Andmekogude tööühma ankeetküsitlusest selgus, et koosvõime tagamisel on kesksel kohal andmekvaliteedi küsimus (vastused küsimusele nr 6), vt Lisa 10. Võib väita, et loenduse tegemisel andmekvaliteedi näitajad mõjutavad infosüsteemi kvaliteeti ja vastupidid, vt Joonist 3.

Registripõhiseks loenduseks on otstarbekas tekitada infosüsteemide võimekus andmete vahetamiseks. Siinkohal tuleks rääkida süsteemist K. Rava omistatud tähenduses. Selle käsitluse järgi on süsteem ühiste eesmärkide nimel töötavate objektide kogum, mis omab infot operatsioonide sisendite ja väljundite kohta ning süsteemide omavaheline suhtlus toimub läbi liidese (Rava, 2000).



Joonis 3. Infosüsteemi ja andmete olulisemad kvaliteedi näitajad koosvõimeks

Seega on andmete liigutamise kontekstis oluline tagada andmekvaliteet ajas: ajakohasus, loendushetkele ajakohastatud õiged andmed. Andmete tähenduse kvaliteet peab olema asjakohane ja usaldusväärne ning andmed peavad olema sobivala kujul. Rava järgi tuleks tööd jagada kahte kategooriasse: infotööd andmetega ja süsteemi tööd (*Ibid*). Selgitasin intervjuude käigus välja takistavad tegurid stabiilse andmeedastuse loomisel, nendeks on:

1. infosüsteemide liidestamise dokumentatsioon pole olnud piisavalt detailne;
2. andmekoosseisude kirjeldused on väga töömahukad sh nende kooskõlastamise kestus;
3. edastatakse registri andmebaasi väljavõtteid, mis nõuab info hilisemat selekteerimist vastavalt vajadustele;
4. muudatuste (uuenduste) hõive vajab pidevat monitoorimist ja manuaalset sekkumist;
5. tehnilisel tasandil on esinenud mahu piiranguid üle x-tee andmevahetuskihi andmete edastamisel.

Järeldan, et olukorra parendamiseks tuleb organisatoorsel tasandil panustada rohkem infotöösse. Andmeedastaja huvides on omada selget arusaama, millises vormis infot vastuvõtjale vaja esitada. Infointegratsiooniks tuleb parendada andmeedastamist, selleks tuleb

andmeedastus protsess jagada iseseisvateks komponentideks (Infotehnoloogia Sõnastik, Eesti Standard EVS-ISO/IEC 2382). Seejärel iseseisvad komponendid identifitseerida ja analüüsida tootmisvõimekuse aspektist.

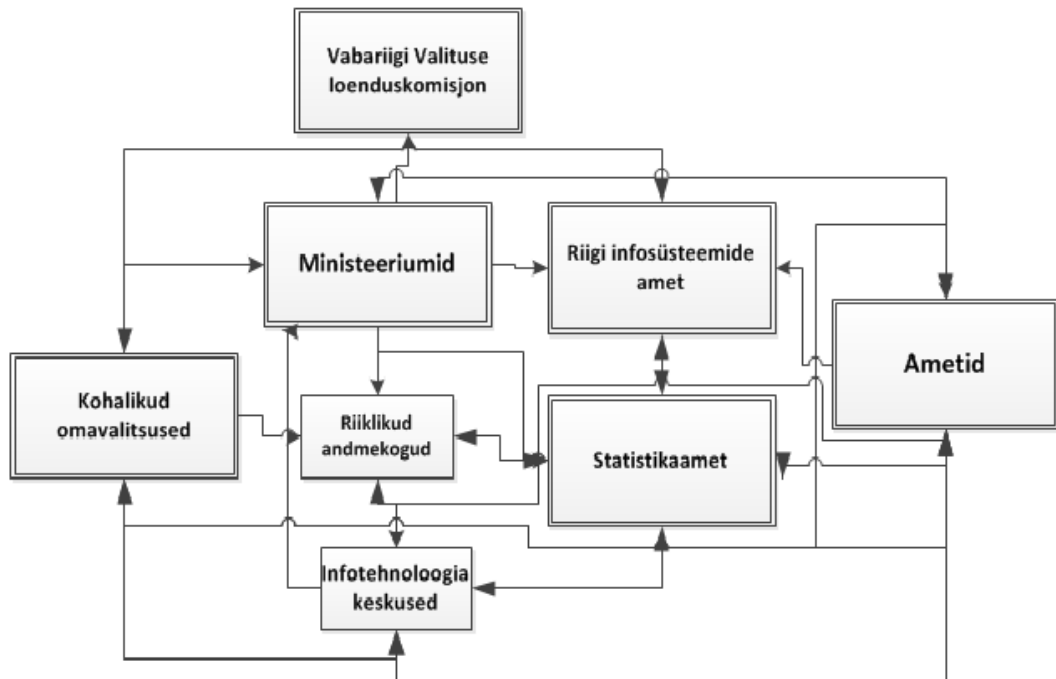
5.1 Infosüsteemide võrgustik loenduseks

Koostööd registripõhiseks loenduseks on vaja teha ministeeriumite vastutusalas olevate andmekogupidajate või infosüsteemi omanikega (VVLK, 16.01.2015):

- Siseministeerium (RR – rahvastikuregister; ETR – elamis- ja töölubade register, hetkel VANGIS tulevikus KIR – riiklik kinnipeetavate, arestialuste ja vahistatute register; KMAIS);
- Haridus- ja Teaduministeerium (EHIS – Eesti Hariduse Infosüsteem);
- Rahandusministeerium (EMTA – maksukohustuslaste register ja TÖR – töötamise register; KPR (KOPIS) – kohustusliku kogumispensioni register, SAP);
- Majandus- ja Kommunikatsiooniministeerium (EHR – riiklik ehtisregister);
- Justiitsministeerium (KR – kinnistusraamat; ARIREG – äriregister; VANGIS(KIR) – riiklik kinnipeetavate ja arestialuste register (tulevik: RKOARR – riigi ja kohaliku omavalitsuse asutuste riiklik register; e-toimik));
- Sotsiaalministeerium (STAR – sotsiaalteenuste ja -toetuste andmeregister; PKR – riiklik pensionikindlustuse register);
- Eesti Haigekassa (KIRST – ravikindlustuse andmekogu);
- Kaitseministeerium (KVKR – kaitseväekohustuslaste register);
- Eesti Töötukassa (EMPIS – töötuna ja tööotsijana arvel olevate isikute ning tööturuteenuste osutamise register);
- Keskkonnaministeerium (Maa-ameti ADS –süsteem);
- Maanteeamet (Liiklusregister).

Alljärgnevalt on esitatud skemaatiline (vt Joonis 4) ülevaade koostöö – ja kommunikatsiooni võrgustikust loenduseks, kus välja toodud institutsioonid, kes on kaasatud loenduse ettevalmistustöösse. Koostööd tuleb organiseerida mitmete asutuste vahel, kriitilise tähenduse omandab siinkohal infoliikumine ja sotsiaalsel tasandil kommunikatiivne integratsioon.

Infosüsteemide koosvõime tagamiseks on asutuste ja organisatsioonide vaheline koosvõime tegutseda võrgustikus (Miles jt, 1992).



Joonis 4. Loenduse korralduse institutsionaalne kommunikatsiooni võrgustik

Infosüsteemide infointegratsiooni probleemid tulenevad registrite loomise ajaloost. Eesti registrid on noored ja need pole loodud loendusstatistika tegemiseks. Registriõhise loenduse tegevussuunised põhinevad paljuski ratsionaalse valiku teoorial. Ratsionaalse valiku fookuses on motivatsiooni eeldused, kus ühtemoodi olulised on väärtused indiviidide tasandil ja eelduste loomine struktuuride tasandil organisatsioonis. Coleman on tuntud kui ratsionaalse valiku süsteemse käitumise selgitajana. Ta on toonud välja, et interaktsioon indiviidide vahel on esile kutsud fenomeni tulemus süsteemide tehnilisel tasandil (Coleman jt, 1992). Koosvõime saavutamiseks on kasulik, et organisatoorsel ja tehnoloogilisel tasandil oleksid süsteemid ühendatavad, mis eeldab standardiseeritust ja avatud tehnoloogiate kasutust, mis võimaldab asutusele tegutseda efektiivsemalt (Porter, 2002). Võrreldes Eesti püüdlusi registripõhiseks loenduseks maailma kogemustega, siis erinevate registrite integreerimine ehk andmete sidumine loendusandmestikuga on meetodina kasutusel Rootsis, Taanis. Selline lähenemine eeldab nagu varem mainitud semantikaalast ja struktuurset koosvõimet, mis on tekkinud Põhjamaades aastakümnete jooksul (Lisa 2), tabel registrite kasutamisele võtmise aastate kohta Taanis, Soomes, Norras, Rootis. Tabelist nähtub, et põhiregistrid võeti neis

riikides kasutusele eelmise sajandi kuuekümnendatel aastatel, statistika tegemiseks olid registrid valmis kümnekond aastat hiljem (Lisa 2). Selleks oli vaja andmefailide sisu ja struktuur kokku leppida registripidajaga ning mestida andmed, mis eeldas mõistete ja definitsioonide harmoniseerimist (ÜRO, 2007: 3-33). Andmete mestimise käigus esineb alati andmete vastuolulisuse probleemi, mida raske lahendada riikides, kus on registrites vastuolulised andmed. Andmete edastajatel ja vastuvõtjal ei ole palju teha olukorra parandamiseks, see nõuaks andmete registreerimise kultuuri muutust ühiskonnas. Registreerimise kultuur on Eestis teistsugune kui seda Põhjamaades, kus see on välja kujunenud sajandite jooksul (Tiit, 2012).

Missugused võiksid olla meetmed infosüsteemide koosvõime tagamiseks Eestis? Vastuseks võiks olla, et andmekogude infosüsteemide väljaarendamine toimuks tsentraalselt. Registritevaheline koostöö toimub, kuid tuleks eristada registrite ja infosüsteemide haldamise ja pidamise probleeme jätkusuutlike lahenduste leidmisel.

Riigi infosüsteemide koosvõime mudel loenduseks seisneb süsteemide võimekuses koos toimida vahetades selleks infot sisu kui ka formaati. Antud võimekus eeldab semantikaalast koosvõimet (Ouksel jt, 1999). Saavutades automaatse andmehõive luuakse infosüsteemidele koostalitlusvõime, mis tagab statistika tegijale strateegiliste eesmärkide saavutamise. Süsteemide koostalitluse loomine on infosüsteemi integreerimise tähtsamaid sihte (Klischevski, 2004).

5.2 Infosüsteemide erinevad eesmärgid ja tasemed

Koosvõime tagamiseks ja riigi IT infrastruktuuri paremaks ärakasutamiseks peavad infosüsteemid rakendama riigi loodud kindlustavaid süsteeme (Avaliku teabe seadus §43):

- Riigi infosüsteemi haldussüsteem (RIHA);
- Infosüsteemide andmevahetuskiht (X-tee);
- Infosüsteemide turvameetmete süsteem (ISKE);
- Klassifikaatorite süsteem;
- Aadressiandmete süsteem (ADS standard);
- Geodeetiline süsteem.

Loenduse kontekstis on kõige olulisemaks riigi infosüsteemi haldussüsteem (RIHA), mis annab ülevaate IT ressurssidest. Koosvõime tagamisel on oluline, et andmekogus olevad klassifikaatorid ja vastavustabelid oleksid kooskõlastatud Statistikaametiga vastavalt Riigi infosüsteemi haldussüsteemi (RIHA) nõuetele (Vabariigi Valitsuse määrus „Riigi infosüsteemi haldussüsteem“, RT I, 04.07.2014 ja Vabariigi Valitsuse 10.01.2007 määrus nr 11 „Klassifikaatorite süsteem“). Teiseks oluliseks infosüsteemiks on aadressiandmete infosüsteem (ADS-i infosüsteem), mis on riigi infosüsteemi kuuluvate registrite tugisüsteem (Vabariigi Valitsuse määrus „Aadressiandmete süsteem“, RT I 2007, 71, 439).

Alates eelmise rahva- ja eluruumide loenduse ettevalmistusperioodist on analüüsitud võimalusi aadressiandmete teenuste ja andmekvaliteedi tagamisel. Selles töös on ette tulnud takistusi, mille ületamiseks on arendatud mitmeid ADS-i infosüsteemi realiseerimise (Kiivet ja Parmakson, 2012). Registripõhiseks loenduseks on tagatud hetkel ADS-i infosüsteemi poolsed tehnilised eeldused. ADS infosüsteemi täiendavaid arendustööd on perspektiivikad, sest ADS-i teenuseid saab kasutada laialdasem tarbijaskond, kuna luuakse rohkem eeldusi tarbijate lisandumisele ning andmekvaliteedi tõusule (ADS käsiraamat, 2013). ADS-i infosüsteemi haldaja on taganud loenduseks erinevatele andmekogu pidajatele vajalikud teenused ning andmekvaliteedi. Rahvastikuregistris kui loenduse põhiregistris on ADS-i andmete kasutamine olulisim. Rahvastikuregistri teenuste arendamist tuleb vaadelda koosmõjus ADS-infosüsteemiga. ADS infosüsteem tagab, et riigi infosüsteem toimiks hästi st aadressi muudatuse korral jõuab info võimalikult kiirelt võimalikult paljudesse registritesse, ilma et peaks selleks eraldi tegevusi tegema (*Ibid*). Samuti on aadress ühendavaks lüliks vähemalt kuue erineva registrite andmete vahel. Loenduse tööd, mis eeldavad erinevate registrite andmete kasutamist ja seostamist, eeldavad erinevate registrite korrektset liidestumist ADS-ga. Aadressiandmete süsteemiga liidestumine ning aadresside andmekvaliteet on seejuures üheks oluliseks eelduseks. Juhtumisi ei saa mööda minna, et aadressiandmete standardi kasutusest riigis. Aadressiandmete kasutus ja kvaliteet paraneb kui on tagatud aadressiandmete kasutamise ühetaolisus. ADS liidestus tagab riigis tsentraliseerituse põhimõttel aadressiandmete andmevoo. Andmed tulevad Maa-ameti Aadressiandmete süsteemist (ADS), mille eesmärk on tagada aadresside ühene identifitseerimine nii nende asukohas kui ka andmekogudes ning muuta võrreldavaks erineval ajal ja eri põhimõtetel esitatud koha-aadressid. Aadressiandmete süsteemiga liitumisel on alati kasutusel ajakohased, täpsed ja üle-eestiliselt kasutatavad aadressid, mida on võimalik

automaatselt seostada teistes registrites ja infosüsteemides nii avaliku kui ka erasektori andmete tarbijate poolt (*Ibid*).

Kolmandaks oluliseks tugisüsteemiks registripõhiseks loenduseks on andmevahetuskiht x-tee, mis lubab turvalist andmevahetust riigi infosüsteemide vahel. X-tee pakub mitmesuguseid lahendusi nt autentimist, autoriseerimist, logide töötlemist ja krüpteeritud andmeliiklust(Infosüsteemide andmevahetuskiht §2-5). Riigi infosüsteemi kuuluvate registrite vahel tuleb andmete edastamiseks kasutada andmevahetuskihti (AvTS § 439 lg 5). Teema loenduse ettevalmistamisel on vähest käsitlust leidnud seoses andmete edastuse ja vahetuse turvalise teenuse loomisega massandmete jaoks.

Nimetatud infosüsteemidega tuleks loenduse registrid liidestada. Liidestustööd on vaja läbi viia äriprotsesside, andmete ja tehnilisel tasandil, kasutades selleks vastavaid meetodikaid (Sogenbits, 2012). Liidestustööde sisu on iga andmekogu pidaja jaoks erinev, infosüsteemide liidestustööd võivad tähendada lisanduvaid arendusi registripidajatele seoses muudatustega:

- kliendirakendus;
- X-tee teenustes;
- Teabesüsteemides.

Kui registripõhise loenduse ettevalmistustööde käigus saavad andmekogud liidestused, kaasneb sellest riigile:

- suurem efektiivsus, seoses andmete seostamisega erinevate infosüsteemide vahel;
- efektiivsem andmete kasutamine;
- manuaalse töö automatiseerimine.

5.3 Riigiinfosüsteemide koosvõime raamistik registripõhiseks loenduseks

Seoseid andmekogusid talletavate infosüsteemide vahel nimetatakse koosvõimeks (Riigi infosüsteemide koosvõime raamistik, 2011). Koosvõime loenduseks seisneb selles, et erinevad infosüsteemid edastaks kokku lepitud formaadis andmeid üle x-tee andmevahetuskeskkonna statistika tegijale. Teiseks, infosüsteemide poolt edastatavad andmed

peavad olema kvaliteetsed. Statistika tegija andmehõive teenus peab olema andmekogupidajatele arusaadav ja kättesaadav.

Riigis on kehtestatud infosüsteemide koosvõime standardid ja juhendid, mille ülesanne on tagada vähemalt avaliku sektori infosüsteemide toimimine. Rakendusliku iseloomuga on järgmised raamistikud (Majandus-ja Kommunikatsiooniministeerium, 2011):

- 1) Avatud võtme infrastruktuur
- 2) Avatud standardid
- 3) Teenuste mudel
- 4) Geoinfosüsteemide raamistik
- 5) Andmebaaside raamistik
- 6) Riigi IT koosvõime raamistik
- 7) Riigi IT arhitektuur
- 8) Semantilise koosvõime raamistik
- 9) Veebide koosvõime raamistik
- 10) Infoturbe raamistik
- 11) Riigi infosüsteemide arendusprotsess
- 12) Vaba tarkvara kasutus
- 13) Dokumendisüsteemide raamistik

Need raamistikud katavad infosüsteemide arhitektuuri, integratsiooni, ühilduvuse, koostoime, konsolideerimise ja adapteeritavuse teemad. Eestis kehtestatud raamistikud on kooskõlas Euroopa raamistikega (*European Interoperability Framework for European public services*), kooskõla saavutatakse läbi mitme dimensiooni: semantika-, organisatsiooni-, õiguse-, tehnilise koosvõime ja poliitilise konteksti (Majandus-ja Kommunikatsiooniministeerium, 2011, 19). Infosüsteemi koosvõime raamistik loenduseks lähtub rakenduslikest dokumentidest, mis selgitavad koosvõime olemusi: Riigi IT koosvõime raamistik; Riigi IT arhitektuur; Semantilise koosvõime raamistik; Veebide raamistik; Tarkvara raamistik; Infoturbe raamistik.

Koosvõime on oluline selleks, et erinevad infosüsteemid areneksid ajas kooskõllaliselt, ei koguks samasid andmeid ning infosüsteemide poolt pakutavad teenused oleksid lihtsad ja kättesaadavad (Riigi infosüsteemide koosvõime raamistik, 2011). Avaliku sektori asutustele on kokkulepitud koosvõime nõuetest kinnipidamine oluline. Registripõhise loenduse koosvõime tagamise kaugem siht on rahvastiku ja eluruumide loendamise seotud avalike teenuste ümberkorraldamine info- ja kommunikatsioonitehnoloogia abil nii, et pikemas perspektiivis oleks saavutatav teenuste osutamise majanduslik tõhusus ning suureneks asutustevaheline koostöö teenuste osutamisel. Selleks on vaja välja arendada registripidajatel ja statistika tegijal teenus andmete edastamiseks üle X-tee, et andmete liikumine toimuks automaatselt ilma andmehalduri käsitööta ning vastuvõttev asutus saaks valideeritud andmed otse enda infosüsteemi. Kujunenud olukord registrite infosüsteemide baasvõimekuste iseloomustamisel ei hõlma ainult IT arendusi vaid ka tööprotsesside arendamist ja protsessidega seotud andmete korrastamisest.

6. REGISTRIPÕHISE LOENDUSE VAJADUSED JA NÕUDED INFOSÜSTEEMIDE KOOSVÕIMELE

6.1 Riigi infosüsteemide loenduseelne seis

Riigis on olemas kesksed infosüsteemid ja registrid, mis võimaldavad andmeid saada iskute, eluruumide kohta (Tiit, 2015). Järgmise sammuna tuleb tagada riigi infosüsteemide koosvõime, soovitavalt testida lahendusi teisel prooviloendusel, mis toimub 2018 aastal. Koosvõime tuleb tagada loendusaastaks 2020 ja edaspidi. Riigi infosüsteemide koosvõime tagamiseks registripõhiseks loenduseks on vaja, et:

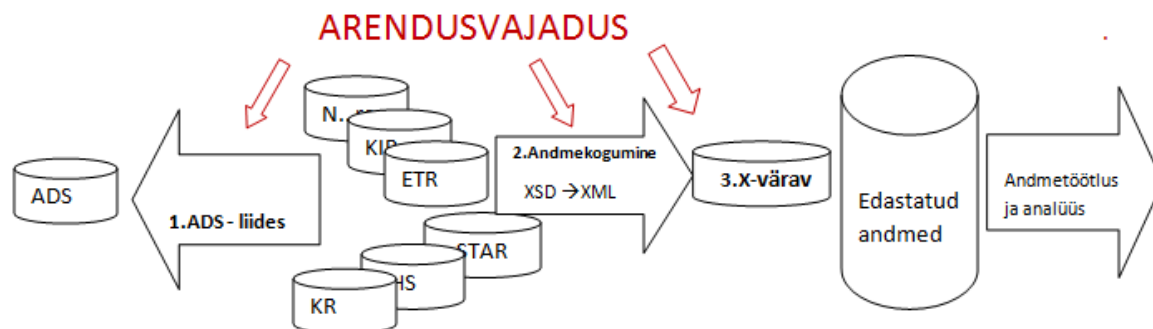
1. kõikides andmekogudes oleval Eesti aadressid, pärineksid Maa-ameti aadressiandmete süsteemi haldussüsteemist;
2. kõik andmekogus olevad isikud peavad olema identifitseeritud isikukoodi (füüsiliste isikute puhul) või äriregistri, mittetulundusühingute ja sihtasutuste registri (äriregistri) koodiga (juriidiliste isikute puhul);
3. tuleb kasutusele võtta andmehõive üle X-tee andmevahetuskihi.

Esimene ja kolmas nõue on riigi spetsiifilised, teine nõue on eeldus registripõhise loenduse tegemisel maailmas (ÜRO, 2007: 7).

Järgmisena selgitasin välja, milline on kõige üldisem registripidajate infosüsteemide koosvõimekuse näitajate seis ADS-süsteemiga liidestamisel ja x-tee andmevahetuskeskkonna kasutamisel andmete edastuseks. Ülevaade on esitatud tabelis (Lisa 7). Ülevaade on koostatud, tuginedes registripõhise loenduse prooviloenduses osalenud andmekogude andmehõive viisi ja ADS liidestuse näitajale. Ülevaade hõlmab 19 registrit ning sellest nähtub, et domineeriv on andmehõives failide vahetus, andmete edastus üle x-tee teenusena Statistikaameti suunal on tagatud neljal andmekogupidajal. ADS-süsteemiga on liitunud või liitumas kokku 8 andmekogu (Lisa 7).

Eeldused üleminekuks registripõhisele loendusele toovad selgelt esile vajaduse riigi infotehnoloogilise sektori kaasajastamiseks. Registripõhise loenduse fookuses on riigi

infosüsteemide arendusvajadus seotud andmekogude liitumisega ADS- infosüsteemiga ning andmete edastuse ja vastuvõtmisega üle x-tee andmevahetuskeskkonna. Vt Joonis 5.



Joonis 5. Riiklike andmekogude infosüsteemide arendusvajadus loenduseks

Allikas: Statistikaamet (autori modifitseeritud versioon)

Andmekogu pidajatele on esitatud registripõhisele loendusele üleminekuks vajalikud tegevused ja nende täitmise tähtsused ning andmekogude kvaliteedi hindamise üldised tingimused, mille täitmise osas seatakse korraldusega andmekogude vastutavatele töötajatele tähtsused (VVLK 16.01.2015). Andmekogude ankeetküsitlusest ilmnes, et liidestustööde täismahulist rakendust piirab ressursi puudus, niisamuti semantilise koosvõime loomiseks vajalik info (Lisa 10). Registripõhise rahva ja eluruumide loenduse ettevalmistus on jõudnud etappi, kus peab kiirendama andmekogude vastutavate ja volitatud töötajate tööd koosvõime tagamiseks, töödemaht on registrite kaupa erinev. Registripõhise loenduse maksumus perioodil 2016 - 2020 on esialgse hinnangu kohaselt kokku umbes 18 miljonit eurot, sellest andmekogude arendusvajadus hõlmab ligikaudu 15% (VVLK, 16.01.2015). Suurem ja kallim tööde maht on seotud registrite andmekorrastusega (*Ibid*). Andmete kvaliteet on registrites tagatud piisaval tasemel väheste tunnuste osas (Lisa 6). Viidatud tööloenduse ettevalmistamisel täidab andmekogude koosvõime ja andmete taaskasutuse eesmärki laiemalt ning loob ulatuslikku institutsionaalse võimekuse Eesti riigile, millest võidavad nii riigi institutsioonid, kohalikud omavalitsused kui ka kodanikud. Kokkuvõtvalt võib öelda, et infosüsteemidega liidestumine eeldab infosüsteemi omaniku suuremat valmisolekut.

6.2 *Strateegiad registripõhise loenduse infosüsteemide koosvõime tagamiseks*

Kui koosvõime on loodud, siis järgmine küsimus kerkib esile seoses valikutega koosvõime tagamiseks. Kontseptsioone levinud infosüsteemide koosvõime strateegiate kohta esitab R.

Klischevski. Eksisteerib kaks nn prominentset strateegiat infosüsteemi koosvõime tagamisel: infointegratsioon ja protsessi integratsioon (Klischevski, 2004). Strateegiate valik on seotud erinevate meetodikate ja tehnikatega. Samas mõlemad strateegiad eeldavad asutuselt üleorganisatsioonilist koostööd partneritega ja administratiivset võimekust. Klischevski väitel on infointegratsiooni eesmärk luua uusi infoteenuseid, mis kasutavad infovoogu jagatud virtuaalses infokeskkonnas (*Ibid*). Protsessi integratsiooni keskmes on protsessi toimimine väljaspool organisatsiooni piire. Protsessi integratsiooni raames loodavad teenused on eelkõige protsesside optimeerimiseks, monitoorimiseks ja kontrolliks, vt Tabel 4. (*Ibid.*).

Tabel 4. Kahe koosvõime strateegia kontseptsiooni võrdlus

Infointegratsioon	Karakteristikud	Protsessi integratsioon
Asutuste vahelised infosüsteemid	Integratsiooni visioon	Asutuste vahelised protsesside võrgustikud
Infoallika disain	Põhistrateegia	Äriprotsesside ümbertegemine
Multilateraalne, globaalne	Partnerluse skoop	Kaasatud vähesed liikmed
Süntaks, semantika, kättesaadavuse kontroll, infokvaliteet	Jagatud abstraktsioonid	Toimimise olemid, protsess ja protsessi kontroll
Madal	Integratsiooni potentsiaal	Kõrge
Madal	Koostöö intensiivsus	Kõrge
Madal	Investeeringute maht	Kõrge

Allikas: *Information Integration or Process Integration*

Infointegratsiooni aktuaalsus uurimisteemana tõstatus 21.sajandi esikümnendil. Uurijate tähelepanu keskmes oli infoallikate erinevad kvalitatiivsed ja kvantitatiivsed omadused ja integratsiooni tehnikad tänases päevas on teema käsitletav laiemalt informatsiooni-kommunikatsiooni teooria osana, kus suurem rõhuasetus on informatsiooni aspektil, mis iseloomustab infoallika ja tarbija suhet (Jhingran jt, 2002). Jhingren kirjeldab infointegratsiooni kui ühte integratsiooni vormi (*Ibid*). Selleks on vaja: struktuurid ühtlustada, infokandjad standardiseerida ja kasutada kokku lepitud kodeeringuid-mustreid. Tuleb

arvestada, et infostruktuur määrab sõnumi tähenduse, viimane sõltub kontekstist ja süsteemi hierarhiast (Parmakson, 2014).

Registripõhise loenduse ettevalmistamisel on fookuses erinevatel etappidel nii üks kui teine strateegia. Kindlasti tuleb infointegratsiooniks teha koostööd erinevate organisatsioonide juhtimise tasandil. Protsessi integratsiooni strateegia võiks järgneda infointegratsiooni strateegiale (Lisa 4), see eeldab vähemate liikmete kaasatust ning on rohkem seotud statistika asutuse strateegilise plaani ühe meetmega kiirendada statistika avaldamist tarbijatele ning vähendada halduskoormust. Esimesel etapil perioodil 2016.-2018. aastatel on soovituslik luua loenduseks infointegreerimise võimekuses, põhjusel, et andmekogumisprotsessis on vaatluse all suurem objektide hulk kui neid on ühes andmeallikas. Soovitatav unikaalandmete seostel põhinev andmehõive eeldab head süsteemide ühilduvust ja standardite kasutamist.

Kas infointegratsiooniprotsessi on võimalik formaliseerida? Väljakutseks on registrite loomise põhjuste mõju arvesse võtmine. Tuleb arvestada, et loenduse puhul on tegemist paljude infoallikatega, mis on pidevas arengus seoses riistvara – ja tarkvaraplatvormide kaasajastamisega. Koosvõime tagamisel loenduseks on oluline sotsio – kultuuriline aspekt protsesside tasandil, see on suutlikkus olla paindlik ja mõista koostööpartneri arendusvajadusi ja võimekusi.

6.2.1 Infointegratsioon

Äriprotsessis tekib kokkuvõtteid, kui andmekogumist ja – korrastamist registriandmetest ei tehta enam failipõhiselt käsitööna ja kui see pole iga andmekogu puhul erilaadne. Registripõhise loenduse suund statistikas eeldab organisatsiooni tasandil keskkonna loomist, mis võimaldaks suuremat tootlust, mitmekesiste dünaamiliste informatsiooni allikate kasutust, juurdepääsu täpsele, relevantsele informatsioonile.

Infointegratsiooni eesmärk on andmetevoog ja tagada juurdepääs struktureeritud informatsioonile, mis on keeruline, sest tootmiskeskond on muutumas komplekssemaks, heterogeensemaks, sest haarab olemasolevaid süsteeme, semantikat, tehnoloogiaid jne. Võtmeküsimus uues olukorras on hallata heterogeensust, vastavat koosvõimet tagades seda semantilisel, süntaksilisel, struktuursel, süsteemi tasemel. (Tabel 5) (Ouksel ja Seth, 1999).

Olukorra teeb keeruliseks asjaolu, et heterogeensust peavad haldama loendustöödega seotud erinevad asutused koos.

Tabel 5. Heterogeensuse olemid infosüsteemides

1. Informatsiooni heterogeensus Semantiline Struktuuriline Süntaktiline jne
2. Süsteemi heterogeensus
3. Informatsioonisüsteemi heterogeensus Andmebaasid Repositooriumid Andmemudelid Juhtimissüsteemid jne
4. Platvormide heterogeensus Operatsioonisüsteemid Riistvara süsteemid Failide süsteemid jne

Allikas: *Semantic Interoperability in Global Information System*

Siinjuures tuleks silmas pidada, et informatsiooni heterogeensusega toimetulekuks peab panustama semantilisele koosvõimele; süsteemi heterogeensus eeldab struktuurset koosvõimekust; informatsioonisüsteemide heterogeensus nõuab süntaktilist koosvõimet ja platvormide heterogeensus süsteemi koosvõimet (Ouksel ja Seth , 1999).

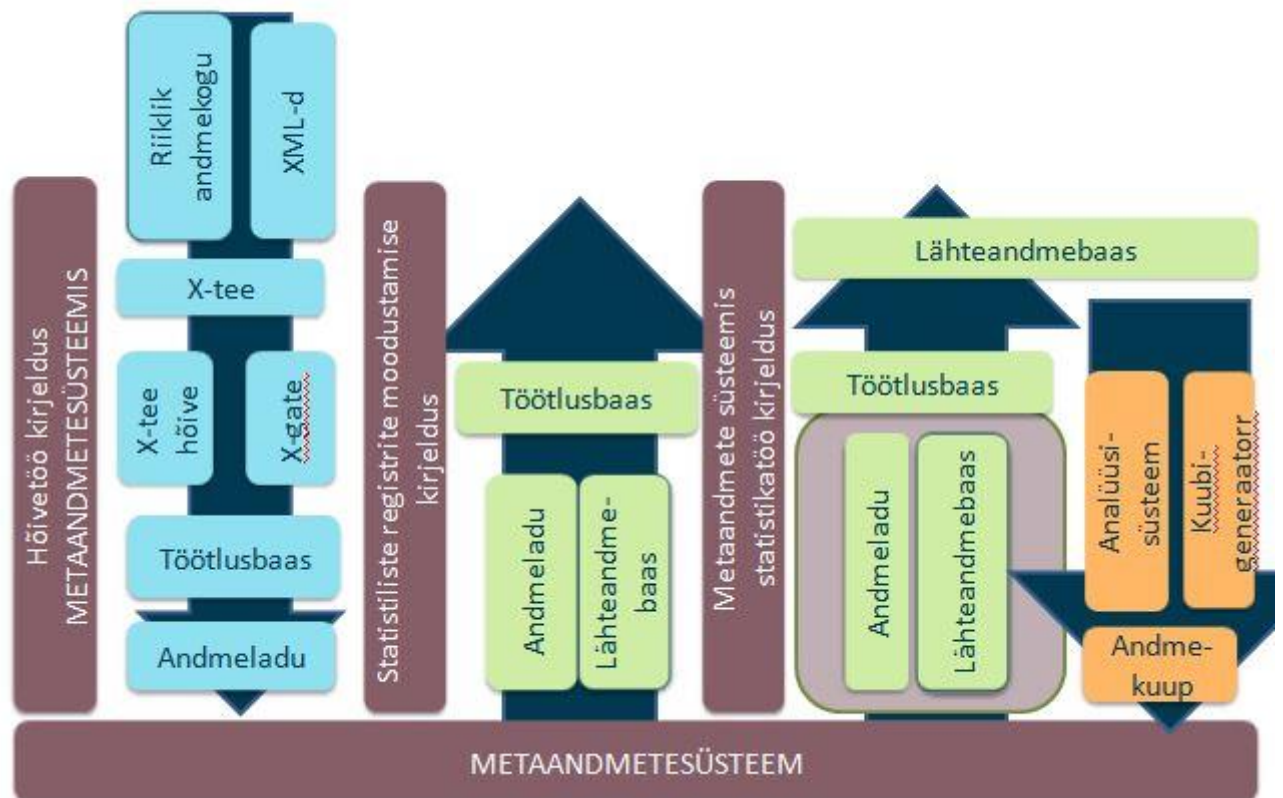
Igasugune info läbib teatud filtrid, mille olemus on otseses sõltuvuses inimese tajust, kogemustest. Süsteemi teooriast (Ludwig von Bertalanffi 1968) lähtudes võiks koosvõime plaanide koostamise kontseptuaalseks aluseks olla printsiipide määratlemine, mis kehtiks süsteemide koostoime kohta üldisemalt (Sistemnye issledovaniya: Ezhegodnik, 1972). Süsteemiteooria seab keskmesse süsteemide ja alasüsteemide vastastikused mõjud koos

tasakaalu saavutamise vajadusega. See teooria on kasutatav uute olukordade analüüsimisel infointegratsioonis ning on võimalik vaadelda süsteemi erinevaid osi neid pidevalt laiendades ja vähendades, leides nii iga tasandi probleemid ja võimalikud lahendused (*Ibid*).

6.2.2 Protsessiintegratsioon

Koosvõime protsessiintegratsiooni keskmes on protsessi jõudlus ja suutlikkus uusi teenuseid juurde luua. Lahendused peavad olema jätkusuutlikud ning toetama jooksva statistika tootmise mahu suurendamist registrite baasilt, optimeerides töömeetodeid, tehnoloogiaid. Eesmärk on toota loendusandmeid vähemalt kaks korda kiiremini kui eelmise loenduse tegemisel st ühe aasta jooksul koguda ja avaldada kõik andmed. Andmehõive eesmärk loenduseks on suurendada automaatse andmehõive osatähtsust järk-järgult. (Lisa 5).

Loendusstatistika tootmise protsess algab loenduse jaoks andmetehõivest. Andmetöötluse käigus korrastatud andmetest moodustatakse lähtebaas. Fikseeritud lähteandmebaasi versioonist genereeritakse väljund. Standardväljund e andmekuup ja lähteandmebaas on aluseks uuele avaldamiseks mõeldud andmestikule (Joonis 6). Registripõhise loenduse infoarhitektuur.



Joonis 6. Registri põhise loenduse infoarhitektuur

Allikas: Statistikaamet (autori modifikatsioon)

Protsessi integratsioon eeldab teenuste juhtimist. Selleks, et vähendada liidestamise ja liidestuste haldamise keerukust, tuleks võtta teenuse pakkujal täiendavalt kasutusele uusi vahendeid, mis aitaksid arhitektuuriliselt eristada integratsioonikeerukuse sisulist äriprotsessi ja funktsionaalsuse keerukust. On otstarbekas jälgida teenuste vahelisi sõnumite suunamisi ja teenuse tõlgendamise võimalusi.

Nõuded infosüsteemide liidestele saavad olla universaalsed:

- kõik infosüsteemi liidesed peavad olema realiseeritud rakenduskihis veebiteenustena;
- kõik pöördumised liideste poole peavad olema turvaliselt autenditud ja autoriseeritud;
- liidesed peavad olema dokumenteeritud vastavalt kehtivale mallile;
- välisliidesed tuleb tarnida koos WADL-i või WSDL-i teenuste kirjeldustega.

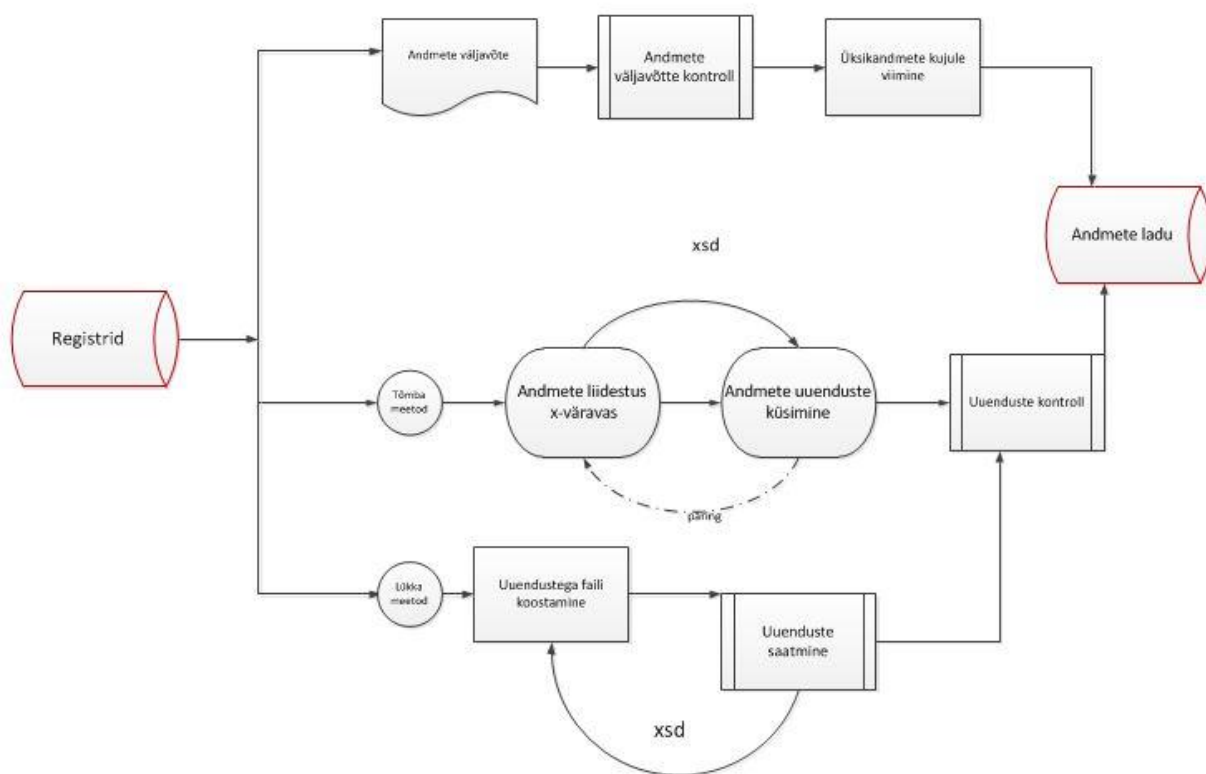
Süsteemide arendamisel oleks hea ärirakenduste ühtsete liidestusvormide ja – tehnikate juurutamine, sest liidestused suurendavad halduskulusid ning tekitavad täiendavat vajadust arendada sisend- ja väljundsüsteeme (Mertins jt, 1998: 589-600).

Andmeedastus võib olla: manuaalne; poolautomaatne, automaatne. Automatiseerimise protsess eeldab, et kõik IT-teenuse osutamiseks vajalikud konfiguratsioonielemendid ja nendevahelised seosed on registreeritud ning läbivad standardiseeritud protsessi, minimiseerides sealjuures võimalikke IT-teenuste katkestusi ja lihtsustades muudatusi (Teenusehaldus ISO/IEC2000-1). Teenuse halduse kasutegur loenduse puhul seisneb selles, et aitab paremini andmehõive muudatusi planeerida, lisaks (Hiie, 2014):

- võimaldab intsidente lahendada kiiremini, mis on oluline tõrgete korral hõives, nt mahu probleemid;
- võimaldab paremini kinni pidada standarditest ja regulatiivsetest kohustustest;

7. ANDMEHÕIVE LOENDUSEKS

Tulles tagasi riigi infosüsteemide koosvõime teema juurde, tuleb see registripõhiseks loenduseks tagada andmetekogumise protsessis, selleks on vaja fikseerida põhimõtteid, mille alusel hakata andmeid registritest hõivama. Luua andmehõive kontseptuaalne raamistik loenduseks. Loenduse jaoks on vaja eesmärgipõhist ühe konkreetse seisuga (loendushetk) andmehõivet. Andmestik formeerub erineval ajal registrites, nt andmed isikute tööhõives olemise kohta, sündide ja surmade kohta jne, seetõttu on vaja tagada väikese mahulist uuenduste hõivet. Automaatse andmehõive kahel meetodil eelduseks (Joonisel 7), loendusandmete hõiveskeem registritest, on universaalne andmeformaad, et loenduseks testitakse andmeid erinevatest andmekogudest, tuleb jätta alles võimalus manuaalseks (andmete väljavõtete) andmeedastuseks.



Joonis 7. Loendusandmete hõiveskeem registritest

Skeem on üldistav ja lihtsustatud ning põhineb andmeedastuse praktikatel, mis selgunud registripõhise loenduse ettevalmistustööde käigus, skeemis on puudu suurandmete hõive.

Selleks on vaja arendada eelnevalt sobiv infrastruktuur ja hõive meetoodika. Andmehõive tehnilisel tasandil tuleb eelnevalt loenduse andmekoosseis iga registri kohta kirjeldada metaandmestikus, kirjeldada registrimuutujaid, mis katavad kogu hõivatava andmestiku. Registrimuutujad tuleb siduda klassifikaatorite, indekse, sõnastikega. Teise sammuna luuakse XSD *lühikese* meetodile, salvestatakse ja omistatakse unikaalne URI. Meetod võimaldab automaatset tagasisidet andmekvaliteedi, st vigade kohta. *Tõmba*-meetodi rakendus oleks vastupidine. (Joonis 7)

Andmete saatmise arhitektuurimustreid on erinevaid. RIA on koostanud kolmeosalisest arhitektuurimudelitest (funktsioon, kontseptsioon, vorm) lähtuva andmete edastusmustril loenduse jaoks, mis ei ole sõltuv andmetest (Kütt, 2014). Edaspidiste tegevuste kavandamiseks on tutvustatud nõudeid neljas kategoorias (VVLK, 16.01.2015):

1. Andmete esmase hõivamise nõue on, et kõik andmed peavad olema esitatud metainfoga sh klassifikaatorite koodidega.

2. Uuenduste hõivamise nõuded:

2.1 Kõik esitatavad andmed peavad olema metainfoga, sh klassifikaatorite koodidega.

2.2 Statistikaameti ja andmekogu vastutava või volitatud töötaja vahelise kokkuleppe kohaselt hõivab Statistikaamet uuendused kas andmekogu vastutava või volitatud töötaja loodud X-tee teenuse kaudu või teeb seda Statistikaameti X-tee teenust kasutades andmekogu vastutav või volitatud töötaja ise.

2.3 Andmed esitatakse XML-vormingus ja andmete kirjelduse esitab X-tee teenuse looja XSD-na. Andmete hõivamisel kasutab Statistikaameti andmekogu vastutav või volitatud töötaja esitatud XSD-vormingus kirjeldust.

2.4 Uuendused peavad sisaldama esitlushetke aega; uuendusi peab saama hõivata alates viimasest esitlushetkest.

3. Nõuded objektide identifitseerimisele

3.1 Kõik andmekogus olevad Eesti aadressid peavad pärinema Maa-ameti aadressiandmete süsteemi haldussüsteemist või olema seotud Maa-ameti ADS-i kirjeldustega (aadressiobjekti identifikaator ADS_OID ja aadressi identifikaator ADR_ID).

3.2 Kõik andmekogus olevad isikud peavad olema identifitseeritud isikukoodi (füüsiliste isikute puhul) või äriregistri, mittetulundusühingute ja sihtasutuste registri (edaspidi äriregistri) koodiga (juriidiliste isikute puhul).

3.3 Nõuded andmekogus olevatele klassifikaatoritele ja vastavustabelitele näevad ette, et need on kooskõlastatud Statistikaametiga vastavalt riigi infosüsteemi haldussüsteemi (RIHA) nõuetele.

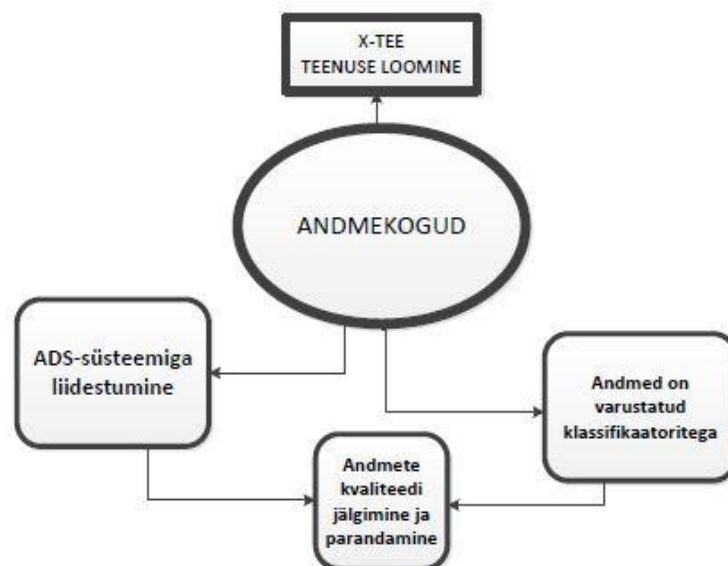
4. Nõuded andmekvaliteedi tagamiseks:

4.1 Andmekogud peavad arvestama Statistikaameti poolt tehtava perioodilise kvaliteedihindamise tulemusi.

4.2 Andmekogudest hõivatavad andmed peavad vastama Statistikaametiga kokkulepitud tunnustele, olema täielikud ja korrektsed.

4.3 Andmekogud peavad kontrollima oma andmete sisemist kooskõla (nt avastada ja parandada topeltkirjeid ja eri tunnuste koosmõjul vigu jmt)

Nõuete täitmiseks vajalikud tegevused kõikide andmekogude puhul on keskselt andmekvaliteedi parandamine ja jälgimine; ADS-i haldussüsteemi andmete kasutuselevõtt muudest allikatest pärinevate aadresside asemel; andmete varustamine klassifikaatoritega ja X-tee teenuse loomine või Statistikaameti X-tee teenuse kasutuselevõtmine (Joonis 8).



Joonis 8 . Andmekogude põhitegevused nõuete täitmisel

Statistika tegija tuleb infosüsteemide loendustegevuse ulatuses tagada:

1. andmekogude koosseisude tehnilise dokumentatsiooni ning rakendama infoturbe abinõud;
2. tehniliste tingimuste tasemel andmekogu vastutava töötleja ja andmekogude volitatud töötlejate vaheliste lepingute sõlmimise;
3. infosüsteemi ja andmekogude vahel koostööd tehnilistes küsimustes riigi kesktasandi infosüsteemidega ning andmekogudega.

Andmehõive nõuete täitmine ja tegevuste elluviimine tagab jätkusuutliku riigi infosüsteemide koosvõime registripõhiseks loenduseks.

7.1 Andme esitusmudel andmehõiveks

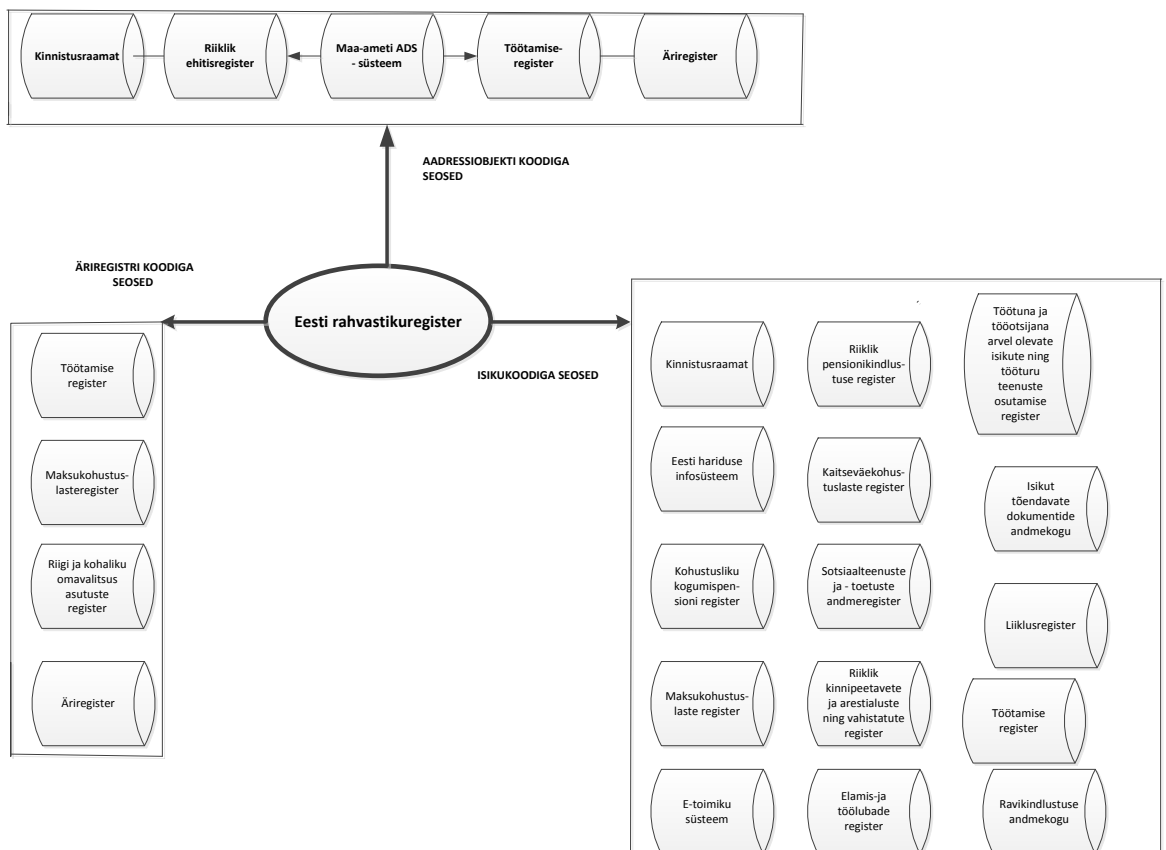
Andmeid, mida loenduse tulemuste saamiseks kogutakse on püütud käsitleda mitmest lähtekohast. Andmeid võib käsitleda infosüsteemi osana ja andmemudel esitab info struktureerimise viisi reaalsest andmeobjektist. Andmemudeleid saab kirjeldada erinevalt nt olem-seosena, loogilisena (UML-d), füüsilisena või reaalsena (Zachmann`i raamistik, 2007). Mudelite loomisel tuleb arvesse võtta nende koostamise eesmärki, äriprotsessi mudelit, infosüsteemi mudelit, tehnoloogia mudelit, füüsilist-ja reaalsel andmemudelit(*Ibid*).

Loendusandmete kogumiseks on otstarbeks loendustunnuste põhjal andmekoosseisud kirjeldada. Andmekoosseisude semantiline kirjeldamine on semantikavarade haldamise töö ja tõrgeteta infoliikumise eeldus andmete edastaja ja vastuvõtja vahel. Semantikaalase koosvõime edendamiseks tuleks eristada erinevaid töö etappe (Küngas, 2011):

- ontoloogia koostamine tunnuste kirjeldamiseks;
- andmekoosseisu koostamine;
- andmekoosseisu semantiline kirjeldamine.

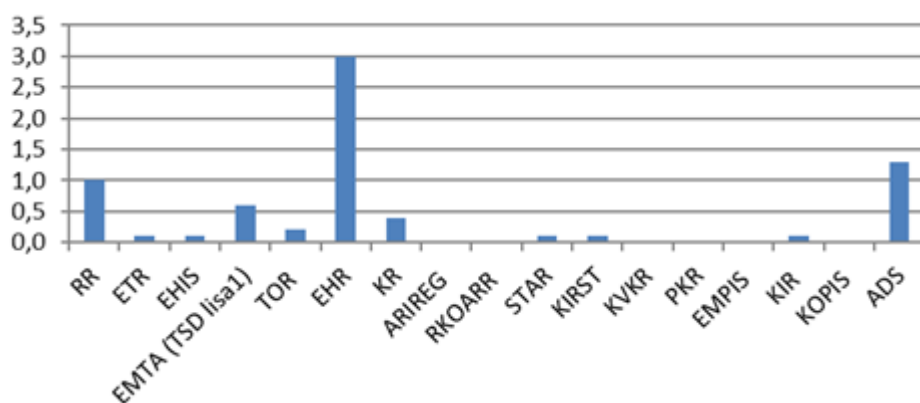
Infosüsteemide koosvõime loomisel on võimalik kasutusele võtta hõive meetod, mis põhineb unikaalsetel andmetel: isikukoodid, äriregistrikoodid, aadressi ID-d. Need on põhiandmed riigi infosüsteemi kuuluvas andmekogus (Avaliku teabe seadus § 43). Unikaalandmetel põhinevas käsitluses üks infosüsteem pakub andmeid kasutada teistele süsteemidele. Unikaalandmete kasutamiseks andmehõives on vaja määrata andmeobjektide omavahelised võimalikud seosed, mis võiksid mingisuguses kontekstis huvi pakkuda.

Kogudes andmeid statistiliste objektide alamhulkade kohta erinevatest registritest ja infosüsteemidest, saab infohulk saab baseeruda mudelil, kus fikseeritakse info tähendus, seosed, formaat jne (Isotamm, 1998). Mudel on alati abstraktne ja selles ei pea kirjeldama kõike, mis on määratud seostega süsteemis (*Ibid*). Koostas in andmehõiveks üldistava andmeseoste mudeli kolmedel unikaalandmetel (aadressiandmed; isikukoodid ja äriregistrikoodid) (Joonis 9). Mudeli koostamisel ei arvestatud piirangutega, mis tulenevad riistvara arhitektuurist või tarkvara kasutuse piirangutest. Kesksel kohal on infoobjektid, unikaalandmed ja nende seostamine registritega. Kolme unikaalse koodiga: isikukood, äriregistrikood ja aadressikood on võimalik seostada 22-e erineva registri andmestikku (andmekogude andmed loenduseks). Edasi tuleb mudelit täpsustada äriprotsesside –, andmete – ja tehnilisel tasandil.



Joonis 9. Unikaalandmetel (äriregistrikoodid, isikukoodid, aadressiobjekti koodid) põhinev andmeseoste mudel.

Mudeli loomise tagamõte oli, et statistika tegemiseks mõeldud andmete kogumine registritest võiks baseeruda otsustel andmete hulga ja mahu kohta, sest loenduste puhul tuleb arhiveerida hõivatavad andmed ehk toorandmed, seejärel dubleerida andmeid andmetöötluse pakettides ja hiljem tagada agregatsioonide väljundtabelid ning lisaks arhiveerida 25. aastaks andmed (EL nr 763/2008). Vaieldamatult suurenevad kulutused serveri mahtudele seoses loendustöödega, kui koguda täiendavalt andmeid, mis olemas koodi sisus (nt sünniaeg ja sugu isikukoodis). Kindlasti on juhtumeid, kus dubleerivate andmete kogumine on vajalik tööks, kuid reeglits võiks saada korrastatud unikaalandmetel (koodidel) andmete kogumine. Alljärgnev graafik Joonisel 10, annab ülevaate ühekordsest andmemahust ca 7GB 17 registri baasil üheks loendustööks. Ilmekas on graafik seetõttu, et näitab umbes 2 korda suuremat andmemahu kui võrrelda kahte kõikset registrit (RR ja EHR). Ühe registri puhul on kasutatav andmete andmebaasi väljavõte andmeedastuseks ja teise registriga toimub andmeedastus üle X-tee andmevahetuskihi.



Joonis 10. Väljavõte: andmehõive maht (GB) SAS-i failide baasil per 17 andmekogu

Üleminek unikaalandmete seostepõhisele andmehõivele eeldab semantikavarade haldust:

- andmekogude andmekoosseisude kirjelduste olemasolu;
- andmekogud peavad omama kirjeldusi RIHA-s;
- registrite andmestik tuleb liigitada andmeallikatena esmajärgulisteks ja teisejärgulisteks.

Käesolev andmehõive esitlusmudel vajab edaspidist arendust. Infoobjekti tasemel on tarvis struktuurid ja vormid ühtlustada ja standardiseerida. Kui on olemas unikaalandmete seostel põhinev kirjeldus registrite kaupa, siis järgmise sammuna tuleb infointegratsiooni käigus viia

kirjeldus vastavusse andmehõive kontseptsiooni infostruktuurile st selgitatakse välja vastavused, seosed ja erinevused infosüsteemide ja infoobjektide vahel.

Esitlusmudeli edasine arendus eeldab süstemaatilist ja järjepidevat tööd andmekoosseisudega, mis osalevad andmehõives. Andmekoosseisud loenduseks ei ole hetkel täpselt fikseeritud, sest registripõhise loenduse meetodika on arendusfaasis. Arendusfaasis tuleb eelkõige integreerida (ühitada) rahvastikustatistika andmed loendusandmetega, see on tarbija huvidest lähtumine. Teiseks registripõhise loenduse meetodika areneb vastavalt Eesti registritesüsteemi arengule, kui on tulemas uusi registreid, on avanemas uued võimalused (Beltadze, 2015).

8. ARUTELU

Selles peatükis annan uurimisülesannetele vastused ja tutvustan empiirilisi tulemusi, mis ei leidnud varasemat käsitlust töös. Magistritöö peamine eesmärk oli tuvastada takistused registripõhise loenduse korraldamisel. Selgus, et infosüsteemide koosvõime registripõhiseks loenduseks eeldab teatud tingimuste täitmist, mis fikseeritud riigi infosüsteemide koosvõime raamistikus 2011. aastast. Üldisemal tasandil on registripõhiste loenduste eeldusteks: andmete olemasolu registrites ja andmeid peab saama seostada. Eestis on välja kujunenud olukord, kus registritega on kaetud peaaegu kõik kohustuslikud loendustunnused v.a amet ja töökoht. On olemas infosüsteemide võrgustik, et loendada isikud, eluruumid. Toimib isikukoodide süsteem ja juurutamisel on aadressiandmete süsteem, astunud on samme X-tee andmevahetuskäsi kasutamiseks andmehõives. Väljakujunenud riigi infosüsteemide eksisteerivas infrastruktuuris on standardite kasutamine eelduseks registripõhisele loendusele üleminekul. Oluline on tagada andmete hõive registritest niisamuti andmete uuendused, kui toimib automaatne andmekogumise protsess, saab rääkida riigi infosüsteemide koosvõimest. Loendusandmete hõive on võimalik ära korraldada paari päeva jooksul, jäävad mõned erandid, mis tingitud sündmuste ajas tagantjärele fikseerimisega registrites nt sünid, surmad. Automaatse andmehõive eelduseks on tehniline keskkond ja andmete formaat ning kvaliteet. Edaspidi on vaja panustada, et tagada koosvõime tööks erinevate registrite andmekoosseisudega. Unikaalandmete seostel põhinev hõive vähendaks koormust andmete vastuvõtmisel ja säilitamisel. Välja tuleb töötada kvaliteedikriteeriumid hõives osalenud andmete hindamiseks (nt tehniliste vigade arv, puuduvad väärtused jn). Kvaliteedi süstemaatiline kontroll andmekogumise etapis soodustab vigade vähendamist registrite andmetes, samas tuleb välja töötada meetodika, kuidas statistika tegija saab kaasa aidata andmekvaliteedi parandamisse registrites ja andmekogupidaja saab registriandmeid korrastada ilma menetlustoimingut rikkumata.

Töös leiab kinnitust hüpotees nr 1. Riigi infosüsteemide koosvõime tekib, kui andmekogude pidajad panustavad võtmeandmete keskele korrashoiule ja võtavad need kasutusele riiklikes andmekogudes. Uuris in klassifikaatori vigade arvu ühes kõige olulisemas registris – rahvastikuregistris kolmel aastal ühe päeva seisuga: 01.01.2014; 01.01.2015; 01.01.2016.

Valisin registri, kus on läbi viidud ADS liidestustööd. Selgus, et rahvastikuregistris peale ADS-süsteemiga liidestumist on klassifikaatori vigade arv pea olematu 2016 aastal.

2014. aastal oli klassifikaatori vigade arv 18 690.

2015. aastal oli klassifikaatori vigade arv 18 354.

2016. aastal oli klassifikaatorite vigade arv 125.

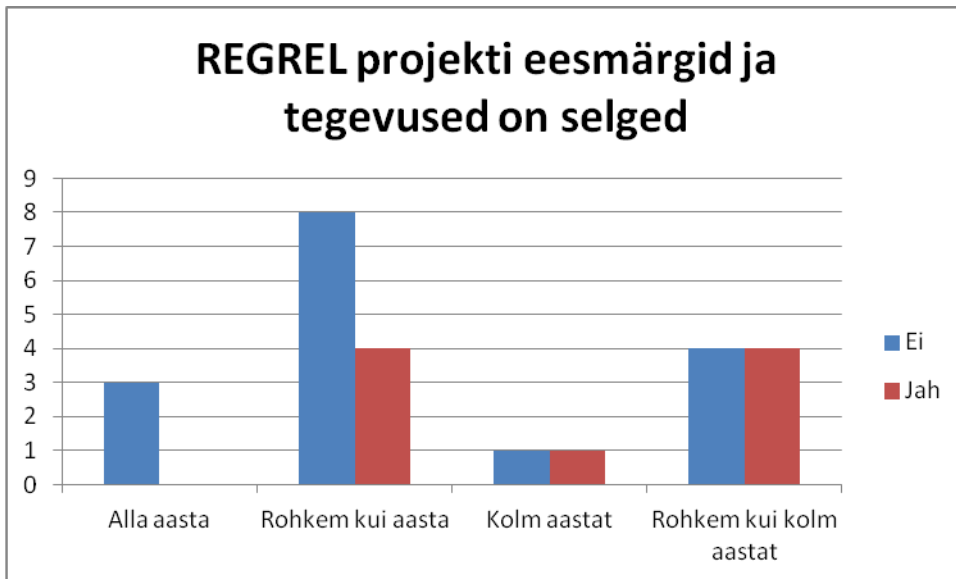
Võib väita, et rahvastikuregistri kvaliteet on paranenud ühe riigi keskse infosüsteemiga liitudes. Rahvastikuregistri ADS-süsteemiga liidestumine võimaldab nüüd paremini isikuid ja eluruumi siduda, see töö on lähtealuseks loendustöödega alustamisel. Kui andmetöötluse käigus tulevad välja ebatäpsed aadressiandmed, saab neid asendada täpsematega varustatud aadressidega. Aadressiandmete täpsustamisel paraneb ka elukoha registreerimise täpsus, nt kui eluruumil on ekslikult mitu aadressi, võivad leibkonnaliikmed olla registreeritud erinevatele aadressidele. ADS-i normaliseeritud aadressi kasutamine parandab tõenäoliselt elukohtade registreerimise täpsust vähemalt 3-5% (Tiit, 2015). On tõenäoline, et peagi teistes andmekogudes ADS-süsteemiga liidestumisel sarnaselt rahvastikuregistriga kvaliteedi näitajad paranevad. Täpseid elukoha andmeid vajab mitte ainult statistika tegija, vaid kogu riik.

Peatun teisel hüpoteesil. Loenduse läbiviimiseks tehtud muudatused riigi infosüsteemides on kasutatavad muude ülesannete lahendamiseks. Tegevused, mis on eelduseks registripõhise loenduse koosvõime tagamisel, ei ole vajalikud ekstra loenduse tegemiseks. Loenduse tegemine peegeldab igas mõttes ainult olukorda. Korras andmeid saab rohkematel eesmärkidel kasutada, kaob vajadus tuletatud andmete jaoks luua eraldi registreid. Näiteks, kui on olemas register, kus hoitakse eluruumidega seotud tunnuseid: eluruumi hoone tüüp, kütte tüüp ja ehitusaasta, on lihtsam tekitada juurde andmestikku küttekolde korrashoiu kohta. Täna eksisteerib selleks kaks lahendust kahe erineva ministeeriumi haldusalas, tekitades registripidamisega seoses halduskoormust. Kui andmed on identifitseeritavad ja kvaliteetsed, on need kasutatavad rohkematel eesmärkidel kui ainult statistika tegemine. Informatsiooni vahetades on võimalik luua uusi teenuseid ühiskonnas või parandada koostööd. Riigi infosüsteemide koosvõime registripõhiseks loenduseks ei ole uus teema. Infoühiskonnas on vaja tagada infosüsteemide koosvõime igapäevaelu korraldamiseks. Uurimistöös tõin välja, et kõik tegevused, mis vajalikud registripõhiseks loenduseks, on loetletud riigi infosüsteemide raamdokumentides ning seotud registrite pidamise ja haldamise temaatikaga laiemalt. Siit tulenevalt leiab osalist kinnitust hüpotees nr 2.

Registripõhiseks loenduseks infosüsteemide koosvõime tagamine võimaldab loendustulemusi saada kiiremini kui tavaloendus. Kui samade registrite põhjal tuleb tulevikus teha loendusi regulaarselt, siis on võimalik kasutada standardiseeritud tarkvara loendustulemuste korrastamiseks ja analüüsimiseks ja kiirendada statistikatootmist tarbija jaoks (Tiit, 2015).

Praegu toimub andmehõive X-tee andmevahetuskihi kaudu ainult nelja registri puhul (Lisa 7). Andmehõivele kuluvat aega saab vähendada, võttes kasutusele standarditel põhineva automaatse andmete edastusviisi ja korraldades ümber registritest andmete kogumise töö ja tehes seda koostöös registripidajatega. Andmehõive töömahukas osa andmekoosseisude täpsustamine vajab reorganiseerimist ning uutel põhimõtetel tööde juurutamist. Soovitav on andmehõives rakendada unikaalandmete seostel põhinevat hõivet.

Koosvõime tagamisel on oluline roll inimfaktoril. Uurisin, mis takistab koosvõime tagamist registripõhiseks loenduseks. Küsimusele vastuse saamiseks viisin läbi registripõhise loenduse andmekogude töörühma liikmete seas ananüümse ankeetküsitluse. Ankeedis oli kokku 7 küsimust (Lisa 9). Küsimustele vastasid kõik andmekogude töörühma liikmed, kes vastutavad registrite eest loenduse ettevalmistamisel. Vastustest selgus vastajate profiil 25-st vastajast oli kolm ja rohkem aastat registripõhise loenduse teemaga seotud 56% vastanutest, alla kolme aasta oli 44%, neist 12% oli teemaga seotud vähem kui aasta (Lisa 10). Vastanutest umbes 2/3 väitis, et mõistab registripõhise loenduse visiooni, kuid pooled nendest ei pidanud oluliseks infot jagada loenduse teemadel oma asutuses, kus nad töötasid. Registripõhise loenduse ettevalmistamiseks loodud projekti eesmärgid ja tegevused ei olnud selged 16-l 25-st. Vt joonis 11. (Lisa 10).



Joonis 11. REGREL projekti eesmärgid ja tegevused on selged.

Paljud andmekogude esindajad väitsid, et koosvõimet registripõhiseks loenduseks ei ole tagatud, sest IT-alaseid nõudeid ei saa täita. Märkimist leidis eelarveliste vahendite nappus arendus – ja/või parendustegevusteks, IT-arenduste paljusus asutustes, konkureerivad lahendused või ei saada lihtsalt aru, mida konkreetselt tuleb teha (Lisa 10). Võib järeldada, et registrite pidamine ja haldustegevused pole piisavalt prioriteetsed ning pole tagatud infopuudusel vajalikke vahendeid, et registreid ajakohastada ja ülal pidada. Ainult 32% andmekogude esindajatest leidis, et on juurutatud standardid ja klassifikaatorid andmekogudes, mis suurendab koosvõimet registripõhiseks loenduseks (*Ibid*). Üldiselt olid vastajatest loendusteemaga vähem aega seotud ametnikud optimistlikumalt häälestatud registripõhise loenduse võimalikkusesse kui rohkemat aega loendusega seotud olnud töötajad (*Ibid*). Pooled vastanutest tõid välja, et koosvõimeks ei ole tagatud õiguslikku raamistikku. Seda teemat tuleb edaspidi uurida, sest uurimistöös esitatud osas registripõhise loenduse õigusraamistiku kohta otseselt ei kinnita seda arusaama. Registripõhiseks loenduseks on olemas õigusraamistik, pigem pole võimalik kõiki tunnuseid koguda (Riikliku statistika seadus paragrahv 22, 2010) registripõhiselt.

Küsitlusest selgus, et koostöövõrgustik koosvõime loomisel registripõhiseks loenduseks on olemas ja asutused panustavad registripõhise loenduse ettevalmistustöödesse. Vastajate keskmine hinnang 10-pallisel skaalal oli 8,4 (Lisa 10). Kõige kriitilisemad probleemid

registripõhise loenduse ettevalmistamisel on andmekogude esindajate hinnangul järgmised(*Ibid*):

1. Andmete kvaliteet põhiregistrites
2. Koostöö puudulikkus registrite pidajate vahel
3. IT-alaste ühtsete standardite ja klassifikaatorite puudumine andmekogudes
4. Konkureerivad IT-lahendused registrite uuendamisel
5. Teenuste kasutuselevõtu tempo
6. Info liikumine
7. Vähene andmekvaliteedi analüüs ja tagasiside andmekogupidajatele
8. Aeg, mis kulub andmekoosseisude täpsustamiseks ja kokku leppimiseks.

Eestis tuleb ettevalmistada registripõhine loendus, mille käigus hõivatakse andmed erinevatest andmekogudest vastavalt andmekaitse ja statistilise turvalisuse reeglitele. Välja töötatud nõuded andmekogudele on piisavad riigi erinevate infosüsteemide koosvõime tagamiseks:

- Kui kõik esitatavad andmed on annoteeritud metainfoga, sh klassifikaatorite koodidega.
- Kui hõive ja andmete uuendused toimuvad üle X-tee teenuse kaudu.
- Kui andmed esitatakse XML-vormingus ja andmete kirjelduse esitab X-tee teenuse looja XSD-na ja uuendused peavad sisaldama esitlushetke aega.

Magistritöös selgitasin välja, missugused on eeldused infosüsteemide koosvõimeks registripõhise loenduse puhul. Infosüsteemide koosvõime mudel loenduseks koosneb infosüsteemide teenuste täiendamisest, kasutajaliideste loomises ning suuremahuliste andmehulkade haldamises. Selgitasin välja, missugust infosüsteemide koosvõime strateegiat kasutada registripõhiseks loenduseks. Esiteks on vaja saavutada infointegratsioon ja seejärel panustada protsessi- integratsioon tegevustesse. Hindasin loenduseks vajalike andmekogude andmehõive toimimise mehhanismide efektiivsust. Andmehõive kõikidest loenduseks vajalikest registritest on ajakulukas, eesmärk on kõik registrite andmed loenduseks hõivata automaatselt. Koostas andmehõive parendamise esitlusmudeli.

Koostasin ettepanekud, kuidas optimeerida andmehõivet ja edendada koostööd registripidajatega.

KOKKUVÕTE

Magistritöös on kirjeldatud riigi infosüsteemide koosvõime tagamise lähtekohti esimese registripõhise loenduse korraldamisel Eestis. Töös on tutvustatud loenduse teemaga seotult koosvõime loomise ja tagamise problemaatikat ajas. Esitatud on infosüsteemide koosvõime olemus loenduseks ja selle tagamise põhimõtted.

Töös on kasutatud avalikku ja ametkondliku dokumentatsiooni analüüsi probleemi, vajaduste ja nõuete analüüsiks. Töös on kasutatud andmehõive spetsialisti intervjuusid andmehõive protsessi hindamiseks. Infosüsteemide koosvõime takistuste välja selgitamiseks viidi läbi ankeetküsitlus registripõhise loenduse andmekogude pidajate seas. Lisaks uuriti kvantitatiivseid näitajaid hõive mahu ja klassifikaatori vigade arvu kohta registris.

Kogutud info põhjal võib järeldada, et tuleb tõhustada koostööd registrite pidajatega. Olukorda on vaja muuta kiiremini, sest loendus peab toimuma nelja aasta pärast. Selleks on vaja eelkõige head selgitustööd, koostööd, võrgustikutööd.

Võttes arvesse Statistikaameti strateegilised kaalutlused andmehõive meetodite kohta, tuleb reorganiseerida andmete kogumine registritest, teist võimalust ei ole andmekogumise protsessi efektiivsemaks muutmisel ja kulude vähendamiseks.

Magistritöös leidsid lahendust ülesanded ja sai täidetud eesmärk. Infosüsteemide koosvõimet luues ja tagades kaasneb tegevus, mis soodustab Eesti infoühiskonna arengut, sest registripõhiseks loenduseks korrastatud ja ühitatud registrite süsteem on suur väärtus, mis võimaldab oluliselt vähendada loenduse kulutusi.

Riigi infosüsteemide koosvõime näitaja on automaatne andmehõive loenduse andmekogumise protsessis. Andmehõiveks on vaja luua stabiilne teenus andmete edastamiseks ja vastuvõtmiseks. Andmehõive on optimaalne, kui see toimub unikaalandmetel.

ETTEPANEKUD

Ettepanek 1

Tuleks kaaluda registritele kesksete kvaliteedinõuete kehtestamist ning üldise andmekvaliteedi kontrolli süsteemi rakendamist riigis. Selleks tuleb töötada välja meetodika. Praegu tegutseb iga register omaette, kuid tulemused oleks ilmselt paremad, kui kogu registrite süsteemi tegevus oleks koordineeritud antud tööloigul.

Ettepanek 2

Registrites olemasolevatele andmetele tuleb rakendada süstemaatiliselt loogilisuse kontrole ja vastuolulised andmed parandada. Vastuoluliste andmete sisestamine ei tohiks olla võimalik.

Ettepanek 4

Vaja on kehtestada riigis üks ja ühtne aadressistandard nii avalikus kui erasektoris.

Ettepanek 5

Ajakohastada registripõhiseks loenduseks Riikliku statistika seaduse loenduse peatükk .

Ettepanek 6

Planeerida ressursse semnatikavarade haldamiseks.

RESÜMEE

Securing Interoperability of State Information Systems for the First Register-Based Census in Estonia

Diana Beltadze

Summary

Present Master`s Thesis deals with the topic of a register-based census and its necessity, state information systems and their present status of networking, the task of securing interoperability and its solution method and, finally, the key role of a statistics collection and processing agency, while securing this interoperability and reaching the necessary result.

The aim of this Master`s Thesis is to evaluate the problems connected with the interoperability of information systems for the register-based census and make proposals for development activities and the improvement of census work processes.

The study focuses on the process of interoperability of public sector organisations` information systems while organizing the first register-based census in Estonia. The aim is to determine the factors hindering the securement of the interoperability of information systems for the register-based census and find strategies helping to solve methodically the situation. There is an implication to stress the importance of the data quality of databases in securing the interoperability of information systems and its topicality.

The arrangement of the first register-based census in Estonia requires a relatively high volume of preparatory work for securing the interoperability of information systems therefore not much of the resources could be economized. From the point of view of horizontal interoperability more vital for the register-based census are the use of standards, improvement of data quality, advancement of semantical capacity. The precondition for securing the interoperability of information systems is the use of state-made securing systems in the information systems (AvTS):

- Management System of the Public Information System (RIHA);
- Information systems data exchange layer (X-Road);
- Information systems security measures system (ISKE);
- Classification system;

- Address data system (ADS standard);
- Geodesic system

The research theme has practical value for the use of register data for statistical purposes.

This is an application-creating research, which maps institutional knowledge and skills necessary for creating a result satisfying certain functional requirements while arranging a register-based census in Estonia in 2020.

This Master's Thesis consists of eight chapters. The first chapter gives an overview of the literature on the interoperability of information systems. The second chapter gives the definition and essence of a register-based census. It was necessary to place it at the beginning of the work because it is closely related to the aim of this work. The advantages of a register-based census are presented. In the third chapter statistics maker's role in implementing a register-based census is introduced. The fourth chapter presents a description of the prehistory of the register-based census in Estonia. The fifth chapter gives an overview of the present status of public information systems for the implementation of the register-based census. An overview of the problems due to different aims and development stages of information systems is presented. The sixth chapter describes the necessities of the register-based census and requirements for the interoperability of information systems. This is the central part of the study – what is actually necessary for securing interoperability? Why is it a problem? The seventh chapter presents a conceptual model of data acquisition. An explanation is given to the requirements databases and statistics makers have to fulfil for this purpose. The eighth chapter includes answers to research tasks and shows whether hypotheses were confirmed.

KASUTATUD ALLIKAD

I Regulaatiivsed andmeallikad

EUROOPA PARLAMENDI JA NÕUKOGU MÄÄRUS (EÜ) nr 223/2009 Euroopa statistika kohta

EUROOPA PARLAMENDI JA NÕUKOGU MÄÄRUS (EÜ) nr 763/2008, 9. juuli 2008, Rahva ja eluruumide loenduste kohta (EMPs kohaldatav tekst), [<http://eur-lex.europa.eu/legal-content/ET/TXT/HTML/?uri=CELEX:32008R0763&rid=1>] 23.03.16

KOMISJONI MÄÄRUS (EÜ) nr 1201/2009, 30. november 2009, millega rakendatakse Euroopa Parlamendi ja nõukogu määrust (EÜ) nr 763/2008, rahva ja eluruumide loenduste kohta seoses andmete ning nende jaotuste tehniliste spetsifikatsioonidega, [<http://eurlex.europa.eu/legalcontent/ET/TXT/HTML/?uri=CELEX:32009R1201&qid=1416579237440&from=ET>] 15.01.16

KOMISJONI MÄÄRUS (EL) nr 1151/2010, 8. detsember 2010, millega rakendatakse Euroopa Parlamendi ja nõukogu määrust (EÜ) nr 763/2008 rahva ja eluruumide loenduste kohta kvaliteediaruandeid käsitleva korra ja struktuuri ning andmete edastamise tehnilise vormingu osas (EMPs kohaldatav tekst), [<http://eur-lex.europa.eu/legal-content/ET/TXT/HTML/?uri=CELEX:32010R1151&qid=1416566627963&from=ET>] 20.02.16

EUROOPA LIIDU ASUTAMISLEPING. Artikkel 285. [<http://www.legaltext.ee/text/-et/T81065.htm>].12.03.16

EUROOPA STATISTIKA TEGEVUSJUHIS. [http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/code_of_practice] 20.02.16

Vabariigi Valitsuse määrus „Aadressiandmete süsteem“, RT I 2007, 71, 439, [<https://www.riigiteataja.ee/akt/12901083?leiaKehtiv>] 12.03.16

Vabariigi Valitsuse määrus „Riigi infosüsteemi haldussüsteem“, RT I, 04.07.2014,
[<https://www.riigiteataja.ee/akt/104072014007?leiaKehtiv>]12.03.16

Vabariigi Valitsuse 10.01.2007 määrus nr 11 „Klassifikaatorite süsteem“
[<https://www.riigiteataja.ee/akt/12910889>] 15.03.16

Vabariigi Valitsuse korraldus.Vabariigi Valitsuse loenduskomisjoni moodustamine. RT
III,12.08.14,6

Riiklik statistika seadus. RT I 2010,41,241

Isikuandmete kaitse seadus.RT I 2007,24,127

Avaliku teabe seadus.RT I 2000, 92597

Rahandusministri käskkiri 6.03.2015.Statistikaameti strateegia 2015-2020

Infosüsteemide andmevahetuskiht §2-5 [<https://www.riigiteataja.ee/akt/12956835>]15.04.16

II Rahvusvaheliste organisatsioonide raportid, soovitused, ülevaated

ÜRO 2008.*United Nations (2008). Principles and Recommendations for Population and Housing Censuses,Revision 2. Statistical Papers Series M, No. 67/Rev.2, ST/ESA/STAT/SER.M/67/Rev.2. United Nations Department of Economic and Social Affairs (UNDESA), Statistics Division, New York.*

ÜRO 2006. *United Nations Economic Commission for Europe (UNECE) (2006). CES Recommendations for the 2010 Censuses of Population and Housing. New York and Geneva: United Nations. [www]http://www.unece.org/fileadmin/DAM/stats/publications/CES_2010_Census_Recommendations_English.pdf*

ÜRO 2007. *United Nations Economic Commission for Europe (UNECE) (2007). Register-based statistics in the Nordic countries. Review of best practices with focus on population and social statistics.*[www]http://www.unece.org/stats/publications/Register_based_statistics_in_Nordic_countries.pdf

ÜRO 2008. *United Nations Economic Commission for Europe (UNECE) (2008). Measuring Population and Housing – Practices of UNECE countries in the 2000 round of censuses. New York and Geneva: United Nations.*

ÜRO 2009. Main Results of the UNECE-UNSD Survey on the 2010 Round of Population and Housing Censuses. Conference of European statisticians. Twelfth Meeting. Economic Commission
For Europe: Geneva

ÜRO.2015. Main Results of the UNECE-UNCD. Survey on the 2020 Round of Population and Housing Censuses. WG-meeting. Geneva

III Metoodika raportid, juhised, analüüsid

Küngas, P.(2011). Semantilise kirjeldamise juhised REGREL-i jaoks. Näiteid REGREL-i metaanalüüsi ja detailanalüüsi tulemuste dokumenteerimiseks SOA Trader OÜ.

Künnapas, K.Lindberg, K. (2014). Õiguslik analüüs: Andmekogude alusregulatsiooni analüüs REGREL projekti teostamiseks.

Tikk, E. (2008). REL2011 õigusanalüüs.

Parre, J. (2013). REGREL õigusanalüüs. Õigusruumi kontentanalüüs

Tikk-Ringas, E. (2010). Õigusruumi kontentanalüüs: Rahva ja eluruumide loenduseks vajalike andmete saadavus ja kvaliteet.

RIA.Kütt, A. (2014). Andmete saatmise arhitektuuri mustrid.

Maa-amet. (2013). ADS käsiraamat.

Majandus ja Kommunikatsiooniministeerium (2011). Riigi Infosüsteemide koosvõime raamistik

[<http://www.riso.ee/et/koosvõime/raamistik>] 20.03.2014

Teenuse haldus: ISO/IEC 20000-1. Hiie, I. (2014). Tallina Ülikooli loengumaterjalid.

IV Ametkondlik dokumnetatsioon

1. Ettepanekud registrite täiendamiseks 15.02.2010

2. SA seisukoht registripõhise prooviloenduse tegevuskava kohta 18.05.2009

3. Ettepanekud ministeeriumidele 18.05.2009

4. Lisa 1 Ettepanekud ministeeriumidele 18.05.2009

5. VVRK protokollid 2008. Rahandusministeerium.

6. VVRK protokollid 2009. Rahandusministeerium.

7. VVRK protokollid 2013. Rahandusministeerium.

8. VVRK protokollid 2010. Rahandusministeerium.

9. VVLK protokollid 2015. Rahandusministeerium.

10. Elektrooniline kiri RE:27.03.06. Ehitisregistri kohta küsimused ja vastused.

11. Memo Registrate korrastamise tegevuskava arutelu 19.01.2009

12. Ettepanekud ministriumidele 19.01.2009
13. Lähteülesanne registripidajatele 15.12.2008
14. Lähteülesanne registripidajatele Lisa 1 15.12.2008
15. Lähteülesanne registripidajatele Lisa 2 15.12.2008
17. Lähteülesanne registripidajatele Lisa 3 15.12.2008
18. Tiit.2006. Registrate analüüsi tulemused.
18. Paut.2008. Registrate analüüsi tulemused .
19. Tiit.2008. Registrate analüüsi tulemused .
20. Tiit.2011. Registrate analüüsi tulemused
21. Majandus- ja taristuministri käskkiri 15.-0213 14.06.2015
22. Memo ümarlaua kohtumisest aadressiandmete teemal. Mai/2008

KASUTATUD KIRJANDUS

Alberts, D.S., Garstka, J.J., Hayes, R.E., Siguari, D.A. (2001). *Understanding Information Age Warfare. CCP, 312*

Beltadze, D. (2015a). *Registripõhine loendus Eestis rahvusvahelises võrdluses*. Konverents: Registrid ja suurandmed, ettekanne.

Beltadze, D. (2015b). *Mis on REGREL?* Statistikaameti teabepäev REGREL teemal, ettekanne.

Bertalanfffi, L. (1972). *Istoria i status obshchy teorii system. Sistemnye issledovaniija: Ezhegodnik*.

Benlian, A. & Hess, T. (2010). *IT standard implementation and business process outcomes: An empirical analysis of XML in the publishing industry*. In *Proceedings of the International Conference on Information Systems*. Atlanta: Association for Information Systems, paper 50

Bhatt, G. & Grover, V. (2005). *Types of information technology capabilities and their role in competitive advantage: An empirical study*. *Journal of Management Information Systems*, 22, 2.

Coleman, J.S. & Fararo, T.J. (1992). *Rational Choice Theory: Advocacy and Critique*. Sage, Newbury Park, Calif.

Chen, M. (2003). *Factors affecting the adoption and diffusion of XML and Web services standards for ebusiness systems*. *International Journal of Human-Computer Studies*, 58, 3.

Chiang, R.H.L., Siau, K. & Hardgrave B. C. (1984). *Systems analysis and design. Techniques, methodologies, approaches and architectures*.

- De Lone, W.H., Mc Lean, E.R. (2003). *The De Lone and Mc Lean model of information Systems success: a ten years update. Journal of Management Information Systems, 19(4)9-30*
- Ducq, Y., Chen, D. & Vallespir, B. (2004). *Interoperability in enterprise modelling: requirements and roadmap. Advanced Engineering Informatics, 18. 193-200*
- Fichman, R.G. & Kemerer, C.F. (1993). *Adoption of software engineering process innovations: The case of object orientation. Sloan Management Review, 7-22.*
- Grant, R. (1996). *Prospering in dynamically-competitive environments: Organizational capability as knowledge integration. Organization Science, 7, 4.*
- Hagedoorn, J. (1993). *Understanding the rationale of strategic technology partnering: Interorganizational modes of cooperation and sectoral differences. Strategic Management Journal, 14, 5.*
- Hanseth, O., Jacucci, E., Grisot, M. & Aanestad, M. (2006). *Reflexive standardization: Side effects and complexity in standard marking. MIS Quarterly, 30, special issue.*
- Hart, P.J. & Saunders, C.S. (1998). *Emerging electronic partnerships: Antecedents and dimensions of EDI use from the suppliers' perspective. Journal of Management Information Systems, 14, 4.*
- Hiie, I. (2014). *IT süsteemide ülalhoid. Loengumaterjal. Tallinna Ülikool.*
- Hirschheim, R., Klein, H.K. (2012). *A Short and Glorious History of the Information Systems Field. A journal of the Association of Information Systems, 13,4, 188-235*
- Huges, G. (2004). *A Guide to Developing Regional Information Society Initiatives. European Regional Information Society Association.*

Infotehnoloogia Sõnastik, Eesti Standard EVS-1801 IEC 2382.

IT juhtimise käsiraamat. (1999). Äripäev: Käsiraamat. Tallinn: Baltic Computer Systems

Isotamm, A., (1998). *Infosüsteemide projekteerimine*. Tartu Ülikooli kirjastus.

Jhingren, A.D. (2002). *Information integration: a research agenda*. IBM systems Journal.

Katz, M. & Shapiro, C. (1994). *Systems competitions and networks effects*. The Journal of Economic Perspectives. Vol, 8, No2 33-115

Kauffman, R.J., McAndrews, J. & Wang, Y.M. (2000). *Opening the "black box" of network externalities in network adoption*. Information Systems Research, 11, 1.

Klein, R., Rai, A. (2009). Interfirm strategic information flows in logistics supply chain relationship. *MIS Quarterly*, 33, 4, 735-762

Klischevski, R. (2004). *Information Integration or Process Integration? How to Achieve Interoperability in Administration*. Electronic Government Volume 3183 of the series

"Lecture Notes in Computer Science. 57-65

Kiivet, H., Parmakson, P. (2012). *Registripõhise rahvaloenduse (REGREL) rakendamise takistused aadressiandmete süsteemi (ADS) võtmes*.

Kittchin, R. (2014). *The data revolution: big data, open data, data infrastructures and their consequences*. SAGE. Publications Ltd.

Laast-Laas, J. (1996). *Registrid ja klassifikaatorid rahvaloenduses*. Automaatkodeerimine. AS

Infoekspert;

Langemets, A. (1999). [www] <http://www.sirp.ee/archive/1990/01.04.99/Sots/sots1-3.html>. (19.september 2014)

- Laihonen, A. (1996). *Register-based population and housing census system in Finland*.
Workshop Census2001: Method and Content. Report. Lisbon. Portugal.
- Legner, C., Wende, K. (2006). Towards an Excellence Framework for Business
Interoperability. 19-th Bled e-Conference eValues. Slovenia. Paper.
- Lorents, P. (2001). *Süsteemse käsitlemise alused*. Tallinn: EBS Print.
- Majandus- ja Kommunikatsiooniministeerium (2014). *Eesti infoühiskonna arengukava
2020 (lõppversioon)*. [www] http://www.riso.ee/et/files/IYA2020_Lisa_1_Moisted.pdf
(12.09.15)
- Maksimovitš G., Romanenko & Samoiljuk A. (2007). *Informatsionnõe sistemõ. Moskva*.
- McGinnes, S., Kapros, E. (2015). Conceptual independence. A design principle for the
construction of adaptive information systems. Information Systems. Vol.47, 33-50.
- Mereste, U. (1987). *Süsteemkäsitus: Süsteemsest mõtlemisviisist majandusnähtuste
käsitlemisel*. Tallinn: Valgus, 264.
- Mertins, K., Jochem, R. (1998). *Handbook on architectures of information systems*. Berlin
Springer, 589-600.
- Mikli, T. (1998). *Sissejuhatus infosüsteemidesse*. Tallinn.
- Miles, R.E. Snow, C. C. (1992). *Causes of failure in network organizations*. California
Management Review, Vol 34. No4. 53-72
- Mintzberg, H. (1979). *Administrative Science Quarterly*. Qualitative Methodology. Vol. 24,
No
4. 582-589

- Ouksel, M., Sheth, A. (1999). Semantic Interoperability in Global Information Systems. A brief introduction to the research area and the special section. *SIGMOD Record*, Vol. 28, No1. 5-12
- Ouksel, A., Iqbal, A. (1999). Ontologies are not the Panacea in Data Integration: A Flexible Coordinator to Mediate Context Construction. *Distributed and Parallel Database*, 7, 1-29
- Panetto, H., Molina, A. (2008). Enterprise Integration and Interoperability in Manufacturing Systems. trends and issues. *Computers in Industry*, 59 (7) 641-646
- Parmakson, P. (2014). *Infokäitluse alused*. Loengumaterjal. Tallinna Ülikool [www] <http://infosysteemid.blogspot.com>. (13.02.16)
- Porter, M. E. (2002). *Strategy and the Internet*. Harvard Business Review.
- Porter, M.E. (1991). *Towards a Dynamic Theory of Strategy*. Strategic Management Journal, 12.
- Praust, V. (1998). *Infoühiskond ja selle teetähised. – Infotehnoloogia haldusjuhtimises*.
- Rava, K. (2000). *Eeldusi infosüsteemide projektide õnnestumiseks*. A & A, nr 1.
- Rits, K. (2010). *Infoühiskond. Information Society*. <http://www.stat.ee/38007>
- Rai, A., Tang, X. (2010). Leveraging IT capabilities and competitive process capabilities for the management of international relationship portfolios. *Information Systems Research*, 21. 3, 516-542

Rai, A., Pavlou, P. A., Im,G.; Du, S. (2012). Interfirm IT capability profiles and communications

for concreting relational value. Evidence from the logistics industry. *MIS Quarterly* 36,

1, 233-262

Sakkeus, L. Puur, A., Aben, S.(2013).REGREL metoodika väljatöötamise raport.Tallinn

Sassen,S. (2002). Towards sociology of information tehnology. *Current Sociology*. CSI. Vol. 50,

No3 1-24

Schulte-Nordholt, E., Hartgers M., Gircour R. (2004). *The Dutch virtual Census 2001. Analysis*

*and Methodology. The Hague: Statistics Netherlands. [www]
<http://www.cbs.nl/nr/rdonlyres/d1716a60-0d13-4281-bed6-3607514888ad/0/b572001.pdf>*

Sheth, A. (1998). *Changing Focus on Interoperability in Information Systems. From System, Syntax, Structure to Semantics, in Interoperating Geographic Information Systems.*

M. F. Goodchild, M. J. Egenhofer, R. Fegeas and C. A. Kottman Kluwer Publishers.

Sheina, M. (2002). *Integration Everywhere*. In: *Computer Business Review*, 10(6).

Shannon, C. E., Weaver, W. (1949). *The mathematical Theory of Communication*. University of

Illinois Press.

Sillaots, M. (2015). Uurimismeetodid. Loengumaterjal. Tallinna Ülikool.

Sogenbits, T. (2012). Süsteemide integratsioon ja liidestamine.[www]

<http://www.ecoful.ee/ivrk/s%FCsteemide%20integratsioon%20ja%20liidestamine,%20loe%20ng%201.pdf> 13.10.2014

Spagnoletti, P., Za, S. (2011). Securing virtual enterprises: organizational requirements and architectural choices. Proceedings on NETs 2011. Kuala-Lumpur.

Stacie, P. (2013). Information Systems Success: The Quest for the Independent Variables. *Journal of Management Information Systems*. Vol 29, issue 4, 7-62

Tallon, P. Ramirez, R. R., Short, J. E. (2013). The Information Artifact in IT Governance, Towards a Theory of Information Governance. *Journal of Management Information Systems*. Vol. 30, issue3, 141-178

Tiit, E.-M. (2009). Ettepanekud ministriumidele ja ametkondadele tegevuste kohta, mis on vajalikud registripõhise loenduse läbiviimiseks 2011. aastal. Tallinn: Statistikaamet.

Tiit, E.-M. (2010). Registrid ja loendused riigi statistikasüsteemis. Registrid ja rahvaloendus. Eesti Statistikaamet teabevihik 20, lk 15–25.

Tiit, E.-M. (2015). *Ettevalmistused registripõhiseks loenduseks*. Statistikaameti kvartalikirj. Nr 3.

Zhao, K. & Xia, M. (2014). *Forming Interoperability Through Interorganizational Systems Standards*. Journal of Management Information Systems /Spring 2014, Vol. 30, No.4

Zhu, K. & Kraemer, K.L. (2005). *Post-adoption variations in usage and value of ebusiness by organizations: Cross-country evidence from the retail industry*. Information Systems Research, 16, 1.

Zhu, K., Kraemer, K.L., Gurbaxani, V. & Xu, S.X. (2005). *Migration to open-standards interorganizational systems: Network effects*.

- Zachman, J. (2007). *Zachman Institute for Framework Advancement*: [www] <http://www.zifa.com/> 12.04.16
- Tanriverdi, H. (2005). *Information technology relatedness, knowledge management capability, and performance of multibusiness firms*. *MIS Quarterly*, 29, 2.
- Vallner, U. (2009). *Registrite koosvõime*. Eesti Statistika Seltsi 21. konverents, ettekanne.
- Venkatesh, V. & Bala, H. (2012). *Adoption and impacts of interorganizational business process standards: Role of partnering synergy*. *Information Systems Research*, 23, 4.
- Walters, B. A. & Tang, Z. (2006). *IT-enabled strategic management*. Idea group publishing Hershey, London, Melbourne, Singapore.
- Wikipedia (2014). *Infoühiskond [19.september 2014]*. <http://Wikipedia.org/wiki/infoühiskond>
- Wigand, R.T., Picot, A., Reichwald, R. (1997). *Information, Organization and Management: Expanding Markets and Corporate Boundries*. John Wiley and Sons, Chichester, England.
- Wilcox, J.B., Howell, R.D. & Breivik, E. (2008). *Questions about formative measurement*. *Journal of Business Research*, 61, 12.
- Wixom, P.H. & Todd, P. A. (2005). *A theoretical integration of user satisfaction and technology acceptance*. *Information Systems Research* 16 (1) 85-102
- Österle, H., Fleisch, E., Alt, R. (2001). *Business Networking: Shaping Collaboration Between Enterprises*. Springer. Berlin.

MÕISTED JA LÜHENDID

Magistritöös on kasutusel põhimõisted „register” ja „andmekogu” sünonüümidenä. Register on riigi, kohaliku omavalitsuse, avalik-õigusliku või eraõigusliku isiku peetav korrastatud andmete kogum, mille pidamisel kasutatakse automatiseeritud andmetöötlust või mida peetakse käsitsi ja korrastatud vormidel, mis võimaldavad andmetega lihtsat tutvumist või nende mehhaanilist töötlemist, sünonüümiks on riiklik andmekogu (Künnapas ja Lindeberg, 2014). Mõiste „infosüsteem” on defineeritud Vabariigi Valitsuse 20.12.2007 määruses nr 252, mille § 3 lg 1 p 4 kohaselt on infosüsteem andmeid töötlev, salvestav või edastav tehniline süsteem koos tema normaalseks talitluseks vajalike vahendite, ressursside ja protsessidega. Probleemne on riikliku andmekogu vastutava ja volitatud töötleva definitsioon, mis andmekogude põhimäärustes sageli ei eristu ja ei selgu kas IKS-ist ja AvTS-ist tulenevad õigused ja kohustused töötlevatele või mis seaduse alusel töötleva määratud on (*Ibid*). Kasutan töös mõisteid „pidaja”, „haldaja” ja „omanik” sünonüümidenä.

Tabel 6. Mõistete ja lühendite kasutus tekstis

Lühend või mõiste	Selgitus
AKI	Andmekaitse Inspeksioon
Andmekogu	Register
Andmekoosseis	Andmekogusse kogutavate andmete kirjelduste nimekiri
Andmemudel	Andmete struktureerimise viis, millega saab kirjeldada reaalse maailma andmeobjekte. Eristatakse erinevaid andmemudelite detailsuse astmeid
Andmevahetusteenus	Andmekogu poolt elektroonilises keskkonnas pakutav teenus
Atribuut	Andmemudeli olemi tunnus
AvTS	Avaliku teabe seadus
CSPA	Common Statistical Production Architecture
ESS	Euroopa Statistikasüsteem
ESMS	Euro-SDMX Metadata Structure
ESQRS	ESS Standard for Quality Reports Structure
GSIM	Generic Statistical Information Model
GSBPM	Generic Statistical Business Process Model

Lühend või mõiste	Selgitus
AKI	Andmekaitse Inspektsioon
Infosüsteemide semantiline koosvõime	Infosüsteemide ja nende poolt toetavate tegevusprotsesside võime vahetada ja mõista vahetatud andmeid ühtemoodi.
IT	Infotehnoloogia
Infosüsteem	Infohalduskeskkond
IKS	Isikuandmete kaitseseadus
Klassifikaator	Riigi või kohaliku omavalitsuse infosüsteemis kasutatav ühetüübilisi objekte teatud tunnuse või tunnuste hulga alusel liigitav ja üheselt identifitseeriv tähistuseeskiri
Klassifikatsioon	Klassifikaator
Koosvõime	Infosüsteemide ja nende poolt toetavate tegevusprotsesside võime vahetada andmeid ja ühiselt kasutada informatsiooni ja teadmisi.
Loendus	Rahva- ja eluruumide loendus
Mõiste	Teadmusüksus, mille moodustab ühene tunnuste kombinatsioon.
Ontoloogia	Valdkonna kirjeldus, mis on masin- ja inimloetavas vormingus.
REGREL	Registripõhine rahva ja eluruumide loendus
RIHA	Riigi infosüsteemi haldussüsteem
SA-WSDL	<i>Semantic Annotations for WSDL and XML Schema</i> , standard veebiteenuste ja andmestruktuuride semantiliseks kirjeldamiseks.
SDMX	Statistical Data and Metadata eXchange
SA	Statistikaamet
Võtmeandmed	Põhiandmed, unikaalsed andmed andmekogudes

Riiklike andmekogude lühendid

RR – rahvastikuregister; EHIS – Eesti Hariduse Infosüsteem; EMTA – maksukohustuslaste register; EHR – riiklik ehitisregister; KR – kinnistusraamat; ARIREG – äriregister; RKOARR – riigi ja kohaliku omavalitsuse asutuste riiklik register; STAR – sotsiaalteenuste ja -toetuste andmeregister; KIRST – ravikindlustuse andmekogu; KVKR – kaitseväekohustuslaste register; PKR – riiklik pensionikindlustuse register; EMPIS (Töötukassa) – töötuna ja töötusajana arvel olevate isikute ning tööturuteenuste osutamise register; VANGIS – riiklik

kinnipeetavate, arestialuste ja vahistatute register; KPR (KOPIS) – kohustusliku kogumispensioni register; TÖR (EMTA) – töötamise register, ETR – elamis- ja töölubade register Maa-ameti ADS –süsteem; SAP – riigi personali- ja palgaarvestuse andmekogu; e-toimik – E-toimiku süsteem; KMAIS – isikut tõendavate dokumentide andmekogu; Liiklusregister.

JOONISED JA TABELID

JOONISED

Joonis 1. Loenduste metoodika vahetamise põhjused maailmas

Allikas: *UNSD.Report on the Results of a Survey on Census Methods used by Countries in the 2010 Census Round*

Joonis 2. Statistika põhiprotsessi mudel ja selle etapid

Allikas: *Statistikaamet*

Joonis 3. Infosüsteemi ja andmete olulisemad kvaliteedi näitajad koosvõimeks

Joonis 4. Loenduse korralduse institutsionaalne kommunikatsiooni võrgustik

Joonis 5. Riiklike andmekogude infosüsteemide arendusvajadus loenduseks

Allikas: *Statistikaamet (autori modifitseeritud versioon)*

Joonis 6. Regstripõhise loenduse infoarhitektuur

Allikas: *Statistikaamet (autori modifikatsioon)*

Joonis 7. Loendusandmete hõiveskeem registritest

Joonis 8. Andmekogude põhitegevused nõuete täitmisel

Joonis 9. Unikaalandmetel (äriregistrikoodid, isikukoodid, aadressiobjekti koodid) põhinev andmeseoste mudel

Joonis 10. Väljavõte: andmehõive maht (GB) SAS-i failide baasil per 17 andmekogu

Joonis 11. REGREL projekti eesmärgid ja tegevused on selged.

TABELID

Tabel 1. Kõige kallim loenduskulu loendusmeetodi järgi maailmas 2010. aasta loendusvoorus:

Allikas: *UNECE 2014.Measuring population and housing.Practices of UNECE countries in the 2010 round of censuses.*

Tabel 2. Kokkuvõte regstripõhise loenduse teemapüstitusest perioodil 1996-2009

Tabel 3. Isiku- ja leibkonna ning eluruumi tunnused registrites

Allikas: *REGREL metoodikaraport 2013 aastal.*

Tabel 4. Kahe koosvõime strateegia kontseptsiooni võrdlus

Allikas: *Information Integration or Process Integration.*

Tabel 5. Heterogeensuse olemid infosüsteemides

Allikas: *Semantic Interoperability in Global Information System*

LISAD

Lisa 1. Loenduse tunnused

1. Isikutunnused		Märkused
1	Alaline elukoht	
2	Sugu	
3	Vanus	
4	Seaduslik perekonnaseis	
5	Sünniriik/-koht	
6	Kodakondsus	
7	Eelmine alaline elukoht ja praegusesse alalise elukohta saabumise kuupäev või alaline elukoht aasta enne loendust	
8	Leibkonnaliikmete omavahelised seosed	
9	Töökoha asukoht	
10	Hetke hõiveseisund	
11	Amet	
12	Majandusharu	
13	Tööalane staatus	
14	Haridustase	
15	Kas on kunagi välismaal elatud ja riiki saabumise aasta (1980 ja hiljem)	
16	Leibkondade valduse staatus	
17	Kogurahvastik	Tuletatud
18	Paikkond	Tuletatud
19	Leibkondlik staatus / Leibkondlik seis	Tuletatud
20	Perekondlik staatus / Perekonnaseis	Tuletatud
21	Tuumperekonna tüüp	Tuletatud
22	Tuumperekonna suurus	Tuletatud
23	Tavaleibkonna tüüp	Tuletatud
24	Tavaleibkonna suurus	Tuletatud
	Rahvus	Riigisisene tunnus
	Emakeel	Riigisisene tunnus

2. Eluaseme andmed		
25	Elamistingimused	
26	Elamispinna tüüp	
27	Elamispinna asukoht	
28	Tavaeluruumide valdusõigus	
29	Omandi liik	
30	Elanike arv	
31	Kasulik põrandapind ja/või eluaseme tubade arv	
32	Veevarustussüsteem	
33	Tualettruum	
34	Vannituba	
35	Kütte tüüp	
36	Eluruumid hoone tüübi järgi	
37	Eluruumid ehitusaja järgi	
38	Eluaseme asustustiheduse standard	Tuletatud

Allikas: Statistikaamet (Aluseks EL regulatsioon 763/2008)

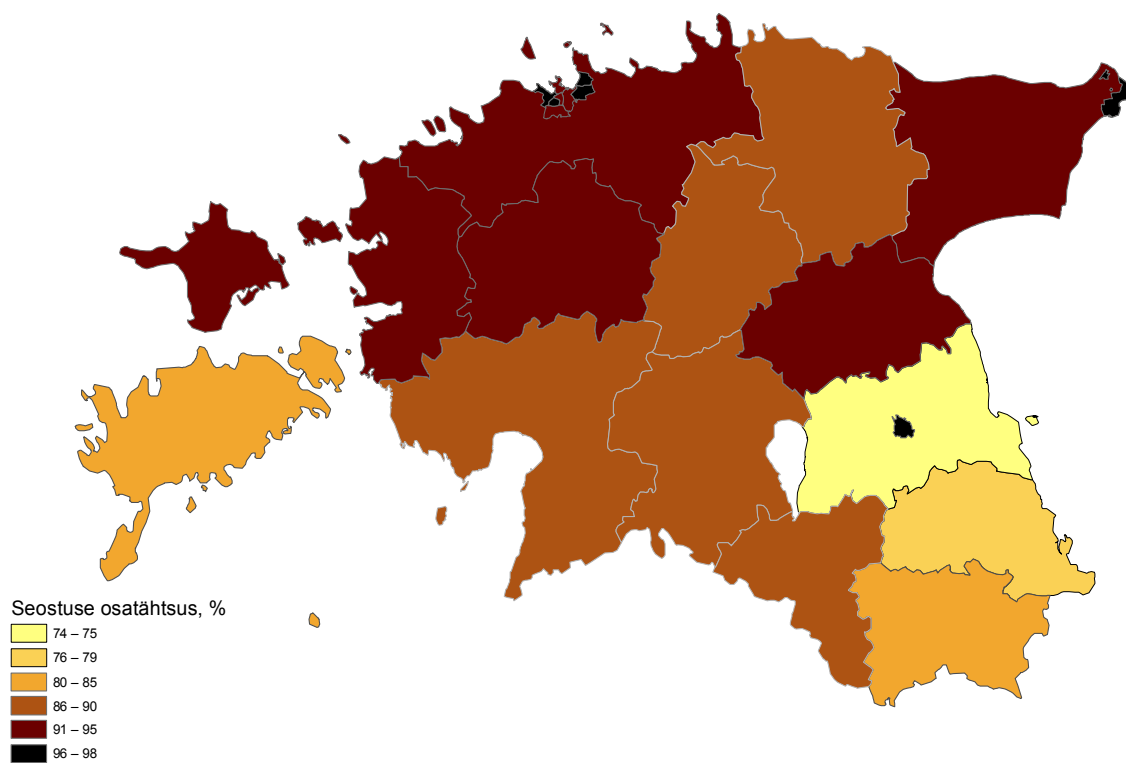
Lisa 2. Registrate kasutusele võtmine loenduseks Põhjamaades

Type of register	Denmark		Finland		Norway		Sweden	
	<i>Established</i>	<i>First used in census</i>	<i>Established</i>	<i>First used in census</i>	<i>Established</i>	<i>First used in census</i>	<i>Established</i>	<i>First used in Census</i>
Central Population Register	1968	1981	1969	1970	1964	1970	1967	1975
Business Register	1975	1981	1975	1980	1965	1980	1963	1975
Dwellings	1977	1981	1980	1985	2001	2011	2008?	2011?
Housing conditions	1977	1981	1980	1985	2001	2011	2008?	2011?
Education	1971	1981	1970	1975	1970	1980	1985	1990
Employment	1979	1981	1987	1990	1978	2001	1985	1985
Family	1968	1981	1978	1980	1964	1980	1960	1975
Household ^a	1968	1981	1970	1975	2001	2011	2011?	2011?
Income	1970	1981	1969	1970	1967	1980	1968	1975
Totally register-based census		1981		1990		2011		2011?

^a Household-dwelling unit, i.e. all the persons living in one dwelling

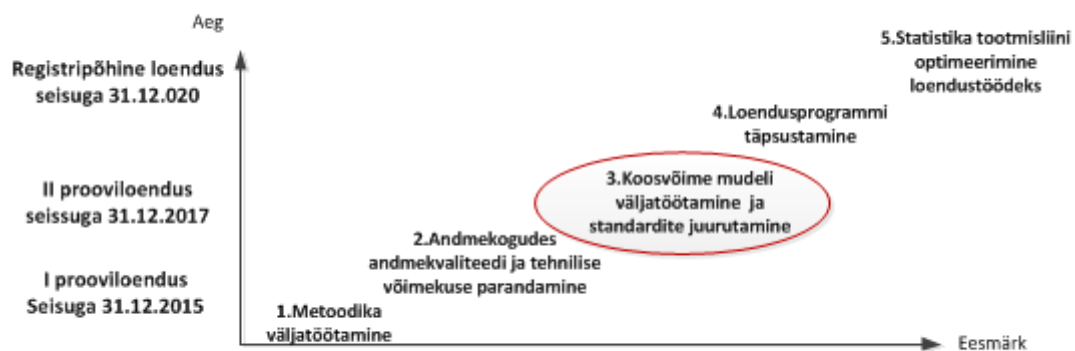
Allikas: ÜRO materjal 2007. lk 5

Lisa 3. Seotud isikud ja aadressid maakondade lõikes 2011. aastal



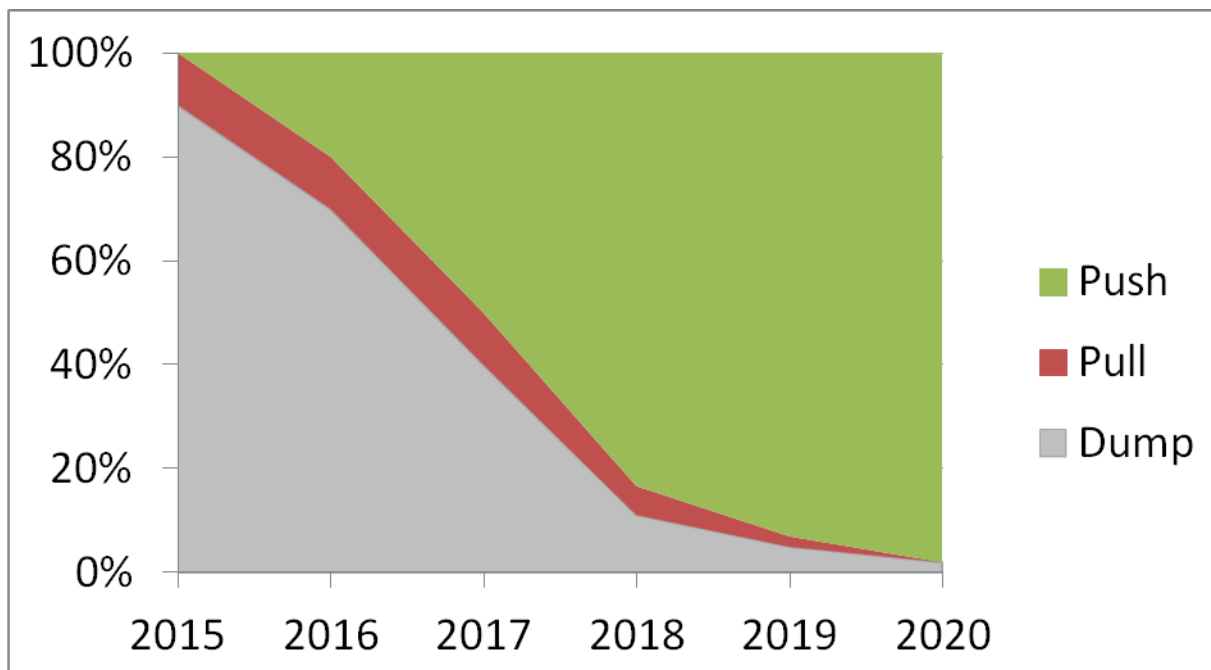
Allikas: Statistikaamet

Lisa 4. Koosvõime tagamise teekaart loenduseks



Autori joonis.

Lisa 5. Loendusandmete andmehõive meetodite strateegia



Allikas: Statistikaamet

Lisa 6. Probleemi kirjeldused loendustunnuste rühmades

Loendustunnuste rühm	Probleem
1.Loendamisele kuuluv rahavastik	Registritele on iseloomulik ülekaetus, mida põhjustab definitsioonide erinevus.
2.Geograafilised ja rändetunnused	Hetkel esineb olukorda, et registreeritud ja tegeliku elukoha andmed erinevad umbes veerandil ühiskonnas.
3.Demograafilised tunnused	Kui elukoht ei vasta tegelikkusele, ei saa partnereid õigesti määrata ja leibkonna struktuur läheb paigast ära (on palju üksikvanemate leibkondi).
4.Sotsiaal-majanduslikud tunnused	Hetkel pole kõikselt hõivatute kohta töökoha ja ameti andmeid.
5.Luibkonna- ja perekonnatunnused	Leibkonnaliikmete omavahelised seosed moodustatakse eluruumi baasil.
6.Eluasemetunnused	Registris esineb ala-ja ülekaetust erinevate tunnuste ja üldkogumi osas.

Allikas: REGREL metoodikaraport 2013. aastal

Lisa 7. Riiklike infosüsteemide koosvõimelisuse seis loenduseks































ANDMEKOGU NIMETUS	X-tee teenus	ADS- liidestus	Märkus	Andmehõive viis
EMPIS	Ei	Ei	-	Serverile fail
ETR	Ei ole	Ei	Klassifikaatorid pole korras	e-mailiga fail
KPR (KOPIS	Ja	Ei	Kasutab pankade aadresse	X-tee
VANGIS (KIR) –	Ei (Tehakse uue registri arenduse pealt (KIR)	Ei	Arendus	Serverile fail
TÖR	Olemas	Olemas		Serverile fail
PKR	Ei	Ei	-	Servile fail
KVKR	Tegemisel	Ja	-	e-mailiga fail
KIRST	Olemas	Ei (pole vaja)	-	Serverile fail
ADS -süsteem	Olemas	Olemas	Jätkuarendus	X-tee
STAR	Ei	Ei	-	e-mailiga fail
RKOARR	Ei	Ei	Liitumine RIK 2016	Serverile fail
ARIREG	Osaliselt	Osaliselt	Arendus	X-tee
KR	Ei	Ei	Arendus	Serverile fail
EHR	Olemas uue	Olemas uue	-	Andmebaasi

	arenduse raames	arenduse raames		väljavõte serverile
EMTA	Olemas	Olemas	Vajab edasiarendust andmekogumise instrumentariumi osas	Serverile fail
EHIS	Tegemisel	Olemas	Klassifikaatorid pole korras	Serverile fail
RR	Olemas	Olemas	Seotud tööd ETR ga.	X-tee
Liiklusregister	Ei	Ei	-	e-mailiga fail
e-toimik	Ei	Ei	-	e-mailiga fail

Autori koostatud tabel.

Lisa 8. Zachmani raamistik

(Adaptation of the Zachman Enterprise Architecture Framework)

	WHAT	HOW	WHERE	WHO	WHEN	WHY
	DATA	FUNCTION	NETWORK	PEOPLE	TIME	MOTIVATION
SCOPE (Contextual)	List of things important to the business  Entity = Class of business things	List of processes the business performs  Process = Class of business process	List of locations in which the business operates  Node = Major business locations	List of organisations important to the business  People = Major business unit	List of event cycles significant to the business  Time = Major Business Event Cycle	List of business goals/strategies  End/Means = Major Business Goal/Strategy
BUSINESS MODEL (Conceptual)	e.g., Semantic Model  Entity = Business Entity Relationship = Business	e.g., Business Process Model  Process = Business IO = Business Resource	e.g., Business Logistics System  Node = Business Location Link = Business Linkage	e.g., Workflow Model  People = Organisation unit Work = Work Product	e.g., Master Schedule  Time = Business Event Cycle = Business Cycle	Business Plan  End = Business Objective Means = Business Strategy
SYSTEM MODEL (Logical)	e.g., Logical Data Model  Entity = Data Entity Relationship = Data Relationship	e.g., Application Architecture  Process = Application Function IO = User Views	e.g., Distributed System Model  Node = KS Function Relationship = Line Characteristics	e.g., Human Interface Architecture  People = Role Work = Deliverable	e.g., Processing Structure  Time = System Event Cycle = Processing Cycle	e.g., Business Rule Model  End = Structural Assertion Means = Action Assertion
TECHNOLOGY MODEL (Physical)	e.g., Physical Data Model  Entity = Segment/Table Relationship = Pointer/Key	e.g., System Design  Process = Computer Function IO = Data Element/object	e.g., Technology Architecture  Node = HW/System s/w Relationship = Line Specifications	e.g., Presentation Architecture  People = User Work = Screen Formats	e.g., Control Structure  Time = Execute Cycle = Component Cycle	e.g., Rule Design  End = Condition Means = Action
DETAILED REPRESENTATIONS (Out-of-context)	e.g., Data Definition  Entity = Field Relationship = Address	e.g., Program  Process = Language Statement IO = Control Block	e.g., Network Architecture  Node = Address Link = Protocol	e.g., Security Architecture  People = Identity Work = Job	e.g., Timing Definition  Time = Interrupt Cycle = Machine Cycle	e.g., Rule Specification  End = Sub-condition Means = stop
FUNCTIONING ENTERPRISE	e.g. DATA	e.g. FUNCTION	e.g. NETWORK	e.g. ORGANISATION	e.g. SCHEDULE	e.g. STRATEGY

Zachmani raamistik mudelite loomisel

Allikas: © Warren Singer 2007, www.technical-communicators.com

Lisa 9. Ankeetküsitluse küsimused

1. Palun märki, mitu aastat oled oma organisatsioonis olnud REGREL teemaga seotud

1.1. Alla aasta

1.2. Rohkem kui aasta

1.3. Kolm aastat

1.4. Rohkem kui kolm aastat

2. Palun hinda info liikumist REGREL andmekogude töörühmas (võib valida mitu vastusevarianti)

2.1. Tean ja mõistan REGREL projekti visiooni

2.2. REGREL projekti eesmärgid ja tegevused on selged

2.3. Olen informeeritud REGREL projektis toimuvast

2.4. Info ei ole arusaadav

3. Palun hinda oma organisatsiooni informeeritust REGREList (võib valida mitu vastusevarianti)

3.1. Tippjuhtkond on teadlik ja tunneb projekti kulgemise vastu huvi

3.2. Infot on jagatud neis osakondades, kus REGREL teemade vastu on huvi ja vajadus

3.3. Jagan infot vastavalt nõudlusele

4. Palun hinda, kas riigis on REGREL korraldamiseks eeldused olemas

4.1. Jah, on olemas

4.2. Ei ole olemas

4.3. Ei oska hinnata

5. Palun märki, millistel tasanditel on riigis infosüsteemide koosvõime REGREL tegemiseks olemas.

5.1. Koosvõime olemas IT-alaste nõuete täitmise tasandil

5.2. Koosvõime olemas õigustasandil

5.3. Koosvõimeks on olemas eelarvelised vahendid

5.4. Koosvõimeks on olemas juurutatud standardid ja klassifikaatorid

5.5. Koosvõimeks on olemas koostöövõrgustik

5.6. Muu (palun nimeta)

5.7. Mitte ühelgi

6. Palun nimeta, mis on REGREL tegemisel probleemsed kohad?

7. Palun hinda 10-pallisel skaalal, mil määral panustab Sinu organisatsioon REGREL õnnestumisse?

(0 – madalaim hinnang ... 10 – kõrgeim hinnang)

Lisa 10. Ankeetküsitluse vastused

IP address	1.Palun märki, mitu aastat oled oma organisatsioonis olnud REGREL teemaga seotud	2.Palun hinda info liikumist REGREL andmekogude töörühmas (võib valida mitu vastusevarianti) [Tean ja mõistan REGREL projekti visiooni]	2.1Palun hinda info liikumist REGREL andmekogude töörühmas (võib valida mitu vastusevarianti) [REGREL projekti eesmärgid ja tegevused on selged]	2.2Palun hinda info liikumist REGREL andmekogude töörühmas (võib valida mitu vastusevarianti) [Olen informeeritud REGREL projektis toimuvast]
213.184.51.117	Rohkem kui kolm aastat	Jah	Jah	Jah
213.184.49.162	Kolm aastat	Jah	Ei	Ei
213.184.51.117	Rohkem kui kolm aastat	Jah	Ei	Jah
213.184.45.205	Rohkem kui aasta	Jah	Jah	Jah
195.80.105.102	Rohkem kui kolm aastat	Jah	Ei	Jah
195.80.104.33	Kolm aastat	Jah	Jah	Jah
193.40.107.131	Rohkem kui kolm aastat	Jah	Jah	Jah
213.184.51.68	Rohkem kui kolm aastat	Jah	Jah	Jah
195.80.105.102	Alla aasta	Ei	Ei	Jah
217.73.4.138	Rohkem kui aasta	Jah	Jah	Jah
213.184.45.205	Alla aasta	Jah	Ei	Ei
213.184.38.214	Rohkem kui aasta	Jah	Ei	Jah
217.73.4.138	Rohkem kui aasta	Jah	Ei	Ei

195.80.105.102	Alla aasta	Ei	Ei	Jah
62.65.237.16	Rohkem kui aasta	Jah	Ei	Ei
193.40.10.2	Rohkem kui kolm aastat	Jah	Ei	Ei
213.184.45.205	Rohkem kui aasta	Jah	Jah	Jah
213.184.45.205	Rohkem kui aasta	Jah	Ei	Ei
195.80.105.102	Rohkem kui kolm aastat	Jah	Ei	Jah
213.184.50.55	Rohkem kui aasta	Jah	Ei	Ei
176.46.62.54	Rohkem kui kolm aastat	Jah	Jah	Ei
213.184.49.162	Rohkem kui aasta	Jah	Ei	Ei
213.184.50.55	Rohkem kui aasta	Jah	Jah	Ei
213.184.38.214	Rohkem kui aasta	Ei	Ei	Jah
213.184.50.55	Rohkem kui aasta	Jah	Ei	Ei

2.3 Palun hinda info liikumist REGREL andmekogude töörühmas (võib valida mitu vastusevarianti) [Info ei ole arusaadav]	3. Palun hinda oma organisatsiooni informeeritust REGREList (võib valida mitu vastusevarianti) [Tippjuhtkond on teadlik ja tunneb projekti kulgemise vastu huvi]	3.1. Palun hinda oma organisatsiooni informeeritust REGREList (võib valida mitu vastusevarianti) [Infot on jagatud neis osakondades, kus REGREL teemade vastu on huvi ja vajadus]	3.2. Palun hinda oma organisatsiooni informeeritust REGREList (võib valida mitu vastusevarianti) [Jagan infot vastavalt nõudlusele]
Ei	Jah	Jah	Jah
Ei	Ei	Jah	Ei
Ei	Jah	Ei	Ei
Ei	Jah	Ei	Jah
Jah	Jah	Jah	Jah
Ei	Jah	Ei	Ei
Ei	Ei	Jah	Jah
Ei	Jah	Jah	Jah
Ei	Ei	Ei	Jah
Ei	Ei	Jah	Ei
Ei	Ei	Jah	Ei
Ei	Ei	Jah	Ei
Ei	Jah	Ei	Ei
Ei	Ei	Jah	Ei
Ei	Jah	Ei	Ei
Ei	Ei	Ei	Jah
Ei	Jah	Jah	Ei
Ei	Ei	Ei	Jah
Ei	Jah	Jah	Jah
Ei	Ei	Jah	Ei
Ei	Jah	Ei	Jah

Ei	Ei	Ei	Jah
Ei	Jah	Jah	Ei
Ei	Ei	Jah	Ei
Ei	Ei	Ei	Jah

4.Palun hinda, kas riigis on REGREL korraldamiseks eeldused olemas	5.Palun märki, millistel tasanditel on riigis infosüsteemide koosvõime tegemiseks [Koosvõime olemas IT-alaste nõuete täitmise tasandil]	5.1.Palun märki, millistel tasanditel on riigis infosüsteemide koosvõime tegemiseks olemas [Koosvõime olemas õigustasandil]	5.2.Palun märki, millistel tasanditel on riigis infosüsteemide koosvõime tegemiseks olemas [Koosvõimeks on olemas eelarvelised vahendid]
Jah, on olemas	Jah	Ei	Jah
Jah, on olemas	Ei	Jah	Ei
Jah, on olemas	Ei	Jah	Ei
Ei oska hinnata	Ei	Jah	Ei
Ei oska hinnata	Ei	Ei	Ei
Jah, on olemas	Jah	Ei	Ei
Ei oska hinnata	Ei	Ei	Ei
Jah, on olemas	Ei	Ei	Jah
Ei ole olemas	Ei	Ei	Ei
Jah, on olemas	Ei	Jah	Ei
Jah, on olemas	Jah	Jah	Jah
Jah, on olemas	Ei	Jah	Jah
Jah, on olemas	Jah	Jah	Ei
Ei oska hinnata	Ei	Ei	Ei

Jah, on olemas	Ei	Jah	Jah
Ei ole olemas	Jah	Jah	Jah
Ei oska hinnata	Ei	Ei	Ei
Jah, on olemas	Jah	Ei	Ei
Jah, on olemas	Ei	Ei	Ei
Ei ole olemas	Jah	Jah	Jah
Jah, on olemas	Ei	Jah	Ei
Ei oska hinnata	Ei	Ei	Ei
Ei oska hinnata	Jah	Ei	Ei
Jah, on olemas	Jah	Ei	Ei
Ei oska hinnata	Jah	Jah	Jah

5.3.Palun märki, millistel tasanditel on riigis infosüsteemide koosvõime REGREL tegemiseks olemas [Koosvõimeks on olemas juurutatud standardid ja klassifikaatorid]	5.4.Palun märki, millistel tasanditel on riigis infosüsteemide koosvõime REGREL tegemiseks olemas [Koosvõimeks on olemas koostöövõrgustik]	5.5.Palun märki, millistel tasanditel on riigis infosüsteemide koosvõime REGREL tegemiseks olemas [Mitte ühelgi]	5.6.Palun märki, millistel tasanditel on riigis infosüsteemide koosvõime REGREL tegemiseks olemas [Teised]
Ei	Ei	Ei	
Jah	Jah	Ei	
Jah	Jah	Ei	
Ei	Ei	Ei	
Ei	Ei	Ei	Toimub regulaarne kohtumine Statistikaametiga, kes loob eeldusi REGREL-i toimumise võimalikkuseks.
Ei	Jah	Ei	

Ei	Jah	Ei	
Jah	Jah	Ei	
Ei	Jah	Ei	
Ei	Jah	Ei	
Ei	Jah	Ei	
Ei	Jah	Ei	
Jah	Jah	Ei	
Jah	Ei	Ei	
Ei	Jah	Ei	
Ei	Ei	Ei	
Ei	Ei	Ei	Vastajal puudub info hindamiseks
Jah	Jah	Ei	
Ei	Jah	Ei	
Jah	Ei	Ei	
Jah	Ei	Ei	
Ei	Jah	Ei	
Ei	Jah	Ei	
Ei	Jah	Ei	
Ei	Ei	Ei	

6.Palun nimeta, mis on REGREL tegemisel probleemsed kohad	7.Palun hinda 10-pallisel skaalal, mil määral panustab Sinu organisatsioon REGREL õnnestumisse-
Kõige kriitilisem ja probleemsem on kvaliteetsete lähteandmete kättesaamine.	8
Lisaks andmekogude poolsetele probleemidele: Statistikaametipoolne projektijuhtimine kipub natuke logisema. Infomaterjalid ei ole ühest kohast kättesaadavad.	9
Põhiregistrite andmekvaliteet. Registreeritud elukoha ja tegeliku elukoha lahknevus. Välismaal elavate, õppivate või töötavate isikute andmed pole aktuaalses seisus.	9
Väga palju erinevaid registreid ja nende vahelise koostöö puudumine ja ühene planeerimine ning ajakavas püsimine.	9
Küsimustele 4 ja 5 on keeruline vastata, sest erinevate riigi infosüsteemide koosvõime hindamiseks peaks tegema teistsugust tööd. Mul ei ole ülevaadet 17 andmebaasi valmisolekust REGRELiks Tööd meie infosüsteemis käivad ja teeme kõik selles, et anda maksimaalne panus REGRELi.	10 – kõrgeim hinnang
Andmekvaliteet ei ole piisav registrites ja seda ei saa ka parandada.	6
Hetkel puudub minul teadmine, kas kõik vajalik info on siiski võimalik registritest saada!	8
Andmete kvaliteet jätab soovida. Näiteks EHR-is.	10 – kõrgeim hinnang
andmed ei ole kvaliteetsed, palju vigu	10 – kõrgeim hinnang
IT-alaste ühtsete standardite ja klassifikaatorite juurutamine andmekogudes, kus nt samu andmeid (nt aadressid) kasutatakse erinevatel eesmärkidel (a'la elukoha	9

registreerimine versus kirja saatmine).	
Registrite andmekvaliteet.	8
Informatsiooni liikumisel Statistikaameti ja registrite vahel on puudujääke. Seda tuleks parandada.	9
Eelarvelised vahendid	8
Konkureerivad IT-arendused	9
Kuna meil on hetkel arendamisel uus register, siis andmete saamine prooviloenduse tarbeks vanast registrist on mõningal määral probleemne (andmed tuleb käsitsi üle kontrollida). Kuid arvan, et pikemas plaanis kaardistame sellega probleemkohad uue registri arendamiseks ning andmekvaliteedi tõstmiseks. Samuti muutub x-tee teenuse tarbeks andmekoosseisu kirjeldus täpsemaks. Seoses andmete ülekontrollimisega on esinenud viivitusi (erinevad osapooled peavad andmed üle kontrollima).	10 – kõrgeim hinnang
Teenuste kasutuselevõtmise tempo. ESA poolt on sisuline andmevajaduse analüüs tagasihoidlik ja andmete tähendust ja konteksti ei täpsustata piisavalt.	7
Kuna minu enda vahetui kokkupuude REGREL teemadega ei ole väga suur, siis võib vastus olla väeinformatiivne. Aga: - erinevate registrite valmisolek andmeedastuseks - uute tehnoloogiliste/tarkvaraliste/riistvaraliste laenduste leidmine ja juurutamine SA-s	8
Ei tea neid	7
Aeg.	9
Ajakava, et tööd kõik tehtud saaksid	7
Varasemalt info liikumine ja semantika, st räägitakse samadest asjadest erinevate nimetustega ja sellest tuleb segadus.	9
Infosüsteemi vaatest: muud arendustööd (sh seadusemuudatusest), mis võivad mõjutada REGREL skoopi	8
Tuleks selgitada rohkem seda, et Kui regrel ei tööta, siis kui palju tuleb riigil raha välja käia.	7

ei oska hetkel midagi nimetada	8
Seni oli probleemseks kohaks eelarvelised vahendid.	8